THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: JOHN M. BAXTER

TITLE OF THESIS: GEOMETRY, RIGID BODIES AND CONSTRAINED MOTION

DEGREE FOR WHICH THESIS WAS PRESENTED: MASTER OF SCIENCE

YEAR THIS DEGREE GRANTED: 1976

THE UNIVERSITY OF ALBERTA

GEOMETRY, RIGID BODIES AND CONSTRAINED MOTION

by

JOHN M. BAXTER ©

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS

EDMONTON, ALBERTA

SPRING, 1976

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled GEOMETRY, RIGID BODIES AND CONSTRAINED MOTION submitted by JOHN M. BAXTER in partial fulfilment of the requirements for the degree of Master of Science.

ABSTRACT

The first part of this thesis is principally a review of a number of geometrical formulations of the classical theory of rigid bodies. Chapter II describes a Lie group model of rotational motion originally due to Arnold [1], and Chapter III incorporates this model into the inhomogeneous symplectic mechanics corresponding to a general Lagrangian and state space. In Chapter IV these ideas are generalized to a homogeneous space-time formulation which includes the relativistic rigid test body of Kunzle [29] as well as a new description of a non-relativistic rigid body based on Galilean relativity.

The second part of the thesis begins with a general discussion on constrained motion and then presents a geometrical theory of constraints. This is applied to a new formulation of constrained motion of a rigid body in curved space-time and several examples are presented to illustrate the methods.

ACKNOWLEDGMENT

It is with great pleasure that I express my gratitude to my supervisor, Dr. Künzle, for suggesting the problems on which the thesis is based, as well as smoothing the progress to their solution. I am also indebted to him for imparting his enthusiasm for Applied Differential Geometry through several courses and informal discussions.

I also extend my appreciation to Olwyn and Georgina for the typing, to Jim Brooke for many stimulating conversations and to Wai-Ling for her constant encouragement.

TABLE OF CONTENTS

PAGE

Abstract iv

Acknowledgment v

CHAPTER I Introduction 1

CHAPTER II Analysis of Rotational Motion of a Rigid Body

§2.1 Introduction 6

§2.2 Review of the Traditional Formulation of the Nonrelativistic Rigid Body 7

§2.3 Notation of Differential Geometry and Lie groups 13

§2.4 The Rotation Group in Three Dimensions 19

§2.5 Canonical Lagrangian Dynamics 25

§2.6 Canonical Dynamics of Rotational Motion of a Rigid Body 30

CHAPTER III A Canonical Lagrangian Formulation for a Nonrelativistic Rigid Body

§3.1 Introduction 33

§3.2 Bundle of Orthonormal Triads as Configuration Space 34

§3.3 Canonical Dynamics of the General Motion 37

CHAPTER IV Motion in Curved Space-Time

§4.1 Introduction 46

§4.2 Galilei and Lorentz Groups and Their Algebras . . 48

§4.3 Galilei and Lorentz Structures on Space-Time . . 50

§4.4 The General Programme of Symplectic Mechanics . . 58

§4.5 Space-Time Model of a Free Rigid Body 62

§4.6 Canonical Distribution on G(V); Motion in Space-Time 69

CHAPTER V Constrained Motion of a Rigid Body

§5.1 Introduction 82

§5.2 Geometrization of the Lagrange Multiplier Formalism 86

§5.3 Geometrical Theory of Constraints 92

§5.4 Application to Constrained Rigid Body Motion . . 95

§5.5 Reformulation of the Problem at a Second-Order System 100

§5.6 Calculation of the Constraint Equations 102

§5.7 Special Cases 107

§5.8 General Solution of the Relativistic Constraint Problem 114

§5.9 Motion Along a Schwarzschild Circular Orbit . . . 121

Footnotes 132

Glossary of Notations 134

Bibliography 138

Appendix 144

CHAPTER I

Introduction

The description of rigid bodies and spinning particles has enjoyed a long and rich history ever since its inception with Euler in the earlier part of the 18th century. Euler's formulation was originally given in the context of Newtonian mechanics; since then, particularly from the end of the 19th century to the present, these systems have been cast into the framework of the various new theories (and formalisms thereof) as they arose. To properly give a historical survey would require the space of a book (indeed, see Corben [8] for a review of spinning particles and gyroscopes in special relativity and quantum mechanics) and is beyond the scope of the present work.

We confine our historical remarks to noting: the emergence of a consistent set of equations of motion for a spinning test particle in a gravitational field with Mathisson [62] in 1937 and later Papapetrou [38] in 1950, the development of a multipole formalism for extended test bodies by Dixon [12], and the existence and uniqueness theorems for a centre of mass of an extended body (Beiglbock [5], Dixon [13], Madore [33]) leading to the equations of motion. All of these and similar derivations proceeded by integrating the conservation law for the stress-energy tensor of an extended body. More recent approaches, such as Souriau [49], Bailey and Israel [3] (charged spinning particles in electromagnetic and gravitational fields), and Hanson and Regge [17] (special relativistic spherical

rigid body), are based on a Lagrangian and the associated variational principle. On a more geometrical level are the canonical symplectic formalisms: the canonical Lagrangian approaches of Künzle ([27], [29]) based on a presymplectic potential 1-form and the canonical Hamiltonian approach of Śniatycki and Tulczyjew [47] based on a symplectic 2-form.

The purpose of this thesis is twofold: to survey the classical formulation of the motion of a rigid body from a geometrical point of view and to develop a geometrical formulation of constrained rigid body motion in both Newtonian mechanics and general relativity.

The contempory interest in rigid test bodies in general relativity seems to be due to, among other things, the gyroscope-satellite experiment that Schiff [43] proposed in 1960 to verify one feature of general relativity that hitherto had been unmeasurable (for example, see [6] and [14] for a discussion of the actual experiment, and [4], [35] and [36] for the theoretical background). This experiment, being carried out by C.W.F. Everitt and W.M. Fairbank and currently near completion, attempts to very precisely measure the precession of the spin axis of a rapidly spinning rigid test body. The various analyses on which the experiment is based are carried out, at least implicitly, in the context of the PPN ("parameterized post-Newtonian") formalism (see [35] for definitions and discussions of this formalism). In this framework, the effect of the earth's rotation upon the gravitational field shows up as off-diagonal terms in the PPN expansion of the metric tensor. It is only in the low-temperature, (almost) force-free environment of the satellite experiment that one hopes to detect this extremely small effect.

Chapter II has the dual function of establishing the notation that will be used throughout as well as presenting a model of rotational motion that will be a key element in the various formalisms to follow. This model is based on the fact that the rotation group SO(3) is a natural configuration space for the orientation of a rigid body.

The next generalization of the formalism is to describe the nonrelativistic rigid body with the aid of the theory of connections on principal fibre bundles; the definitions of this theory are collected in an Appendix. At the same time, Chapter III makes the transition from the Lagrangian formalism to an inhomogeneous symplectic version. Although no new physical content is discerned (or expected!) this modern geometrical description of the (nonrelativistic) rigid body is useful to have (as a point of reference) for comparison purposes with the corresponding Newtonian and relativistic space-time descriptions. Also, development of the relativistic formalism can be guided by analogy with the nonrelativistic version.

Chapter IV is mainly a review of the homogeneous formulation of symplectic mechanics (time is included as a coordinate) and its application to the description of rigid test body motion in a curved space-time. Dynamics is now discussed on a higher level than that of the traditional Lagrangian (or Hamiltonian) in that a (pre) symplectic potential 1-form or presymplectic 2-form will be the basis for dynamics. These two approaches, not necessarily equivalent, have varying degrees of advantage depending on the situation to which they are applied, an example being as simple as the facility of choice of Lagrangian or symplectic form for a particular system.

The material of Chapter IV is used in Chapter V to obtain a symplectic formalism of constrained motion of a rigid body in space-time (the situation envisioned is the centre of mass of the body being constrained to a given worldline). While no new effect is expected in the Newtonian theory, we do expect something in relativity, based on the following analogy: It is known (§4.6) that a relativistic spinning rigid test body freely moving in a curved space-time is subjected to a force which causes the body to deviate from geodesic motion (in a fashion analogous to that of a charged particle in a magnetic field). In analogy to the situation in Newtonian mechanics (holding a top fixed at one point causes it to precess), we expect that, if this body is constrained to (say) a geodesic, this force would be manifested in a precession of the spin axis.

Space limitations make it difficult to define all the mathematical terminology which is used throughout. Only key definitions are given in the text, while a major part of the notation is summarized in the Glossary at the end.

The Einstein summation convention is assumed whenever an index is repeated in an upper and lower position within one term: $x^\alpha x_\alpha \equiv \sum_\alpha x^\alpha x_\alpha$, while the range of an index may be either explicit or implicit; the range for the latter is to be taken from context. Greek indices usually refer to space (-time) components while Latin indices refer to general (triad; tetrad) frame components. Most often capitalized indices $(A,B,\Gamma,\Delta$ for example) assume the values 1 , 2 or 3 while lower case indices $(a,b,\alpha,\beta$, etc.) range over 0 , 1 , 2 , 3 .

Capital Latin indices are always raised and lowered with the Kronecker delta: δ^{AB} or δ_{AB} , and lower case Greek (Latin) indices are raised and lowered with a Lorentz (Minkowski) metric in the relativistic case only.

The organization of the thesis is in the Chapter-Section format. The numbering of equations is by section, so that 1.1(1) refers to equation (1) of §1.1; major definitions are also referred to by number. To facilitate cross-referencing, sections are further divided by lower case Latin letters; for example, 1.1a. The occasional footnote is indicated by a superscript number at the appropriate location.

CHAPTER II

Analysis of Rotational Motion of a Rigid Body

§2.1 Introduction.

Although the subject matter of this chapter is familiar to everyone in its traditional form (eg, see Goldstein [16]), the reformulation of rotational motion in terms of Lie groups may not be so familiar in spite of its inclusion into several of the recent books on mechanics ([42], and [52] for example). Since the end result of this chapter will be the model of rotational motion of a rigid body that we use throughout, we develop in some detail the modern differential-geometric formalism used by Arnold in his now classical work on mechanics on Lie groups [2] as well as subsequent treatments ([22], [29], [32]).

The first section defines a rigid body and states some of the relevent classical results, leading up to Euler's equations. The differential geometry that we use is standard; however, due to the great abundance of notations, §2.2 is devoted to presenting the notation, as well as some of the standard results (without proof) that will be used in the sequel. Useful references would be Kobayshi and Nomizu [24], Warner [59], Godbillon [15] and the surveys found in [19] and [54]. We discuss next the inhomogeneous formulation of canonical dynamics with emphasis on the Lagrangian viewpoint; this is then applied to a description of the rotational motion of a rigid body.

§2.2 Review of the Traditional Formulation of the Nonrelativistic Rigid Body.

(a) Foundations. Like that of a point particle, the concept of a rigid body is a mathematical idealization representing physical reality in an approximate sense. In the domain of Newtonian physics, this approximation is actually quite good so we assume the following definition is valid:

(1) Definition. A rigid body is a distribution of mass in space subject to the constraints that the distances defined by each pair of points contained in the distribution shall be constant in time.

The definition implies that at each moment of time the knowledge of four noncoplanar points of the body determines any other by the constraints. "Space" referred to in (1) is Euclidean 3-dimensional space. From a philosophical viewpoint it may be desirable to begin with the more general "affinely-rigid" body studied by Slawianowski [44]. This is a mass distribution in which the affine relations are constrained to be constant in time, no reference to metrical properties being made (physically, affine rigidity corresponds to the possibility of homogeneous deformations of the body). However, we are mainly interested in rigid bodies as in (1). As we shall see, the definition of rigidity in relativity is much less straightforward and indeed, strictly speaking, can't be done at all.

(b) Kinematical considerations. The remark after (1) implies that fixing a set of three linearly independent vectors at a point in a rigid body characterizes the configuration of that body. Thus, the

complete configuration of a rigid body in space, R^3 is specified by

(i) choosing an orthonormal frame $(\vec{k}_1, \vec{k}_2, \vec{k}_3)$ for R^3 parallely translated to each point of R^3) - the space frame;

(ii) giving a point of the body (eg, centre of mass);

(iii) giving the orientation, relative to the space frame, of an orthonormal frame $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ fixed in the body at the point of (ii).

Formulated mathematically: $x \in R^3$ denotes the origin of the body-fixed frame, $(\vec{e}_A\ ;\ A = 1,2,3)$, and the matrix (e^Γ_A) represents the orientation and is defined by

(2) $$\vec{e}_A = e^\Gamma_A \ \vec{k}_\Gamma \ .$$

It turns out that because of (i) and (iii), the matrix (e^Γ_A) is orthogonal; that is, satisfies:

(3) $$e^\Gamma_A \ \delta_{\Gamma\Lambda} \ e^\Lambda_B = \delta_{AB} \ .$$

This, together with (ii) and (iii) above implies that a rigid body has six degrees of freedom.

In practical situations, one most often represents the rotational degrees of freedom by Euler's angles (see [16] for example) but this approach is of little help in our conceptual framework.

Now, at any instant of time, t , we have seen that the orientation of a rigid body can be specified by an orthogonal matrix (e^Γ_A) . If at one instant, $t = t_o$, the body frame was coincident with the space frame $e^\Gamma_A = e^\Gamma_A\ (t_o) = \delta^\Gamma_A$, as time went on the orientation

changes continuously, to $e^{\Gamma}_{A}(t)$, say. Thus, on a physical basis we do not allow space inversions of the body to occur; hence, the restriction that

(4) $$\det(e^{\Gamma}_{A}) = 1$$

is reasonable.

(5) Definition. The collection of all 3×3 matrices satisfying (3) and (4) forms a group called the special orthogonal group (in three dimensions) and is denoted by $SO(3)$.

Note that Euler's Theorem, that the most general displacement of a rigid body with one point fixed is a rotation about some axis, is just the statement that each $a \in SO(3)$ has 1 as an eigenvalue.

(c) Rotating coordinate systems. We assume that we are given a rigid body held fixed at one point and that its rotational configuration is described as in (b).

Given a vector $\vec{G}$ in R^3 , it is useful to know how to relate the expressions of the time rate of change of $\vec{G}$ in either the space frame or the body-fixed frame. To do this we first give the

(6) Definition. Let $t \to e^{\Gamma}_{A}(t)$ be a motion of the body and for each t , let $(\theta^{A}_{\Gamma}(t))$ denote the inverse to $(e^{\Gamma}_{A}(t))$. Then the vector

(7) $$\omega^{A} \equiv \frac{1}{2}\,\epsilon^{A\,L}_{\;K}\,\theta^{K}_{\Gamma}\,\frac{de^{\Gamma}_{L}}{dt}(t)$$

is the angular velocity of the body at time t .

In terms of (7) then,

$$(8) \qquad \left(\frac{d\vec{G}}{dt}\right)_{\text{space}} = \left(\frac{d\vec{G}}{dt}\right)_{\text{body}} + \vec{\omega} \times \vec{G} \quad .$$

(To show this, all that is necessary is to note that:
$\vec{G} = (G_{\text{body}})^A \vec{e}_A = (G_{\text{body}})^A e_A^\Gamma \vec{k}_\Gamma = (G_{\text{space}})^\Gamma \vec{k}_\Gamma$ and differentiate with respect to time using (7).)

(d) Angular momentum and kinetic energy. As in (c), we assume that the rigid body is held fixed at one point. The physical quantities that we shall now introduce: kinetic energy, moment of inertia and angular momentum, are all defined with respect to this point and, in general, depend on this point. These quantities also depend, in a fundamental way, on the distribution of mass of the body. While a mass distribution can be quite general as in the continuum theories of Noll and Truesdell (see [57] for example; a body is a measure space in this theory), the definitions can be carried through quite generally if we assume the body to be composed of a finite number of point mass particles, $\{m_\alpha : \alpha = 1,\ldots,N\}$ say.

As usual then, we define

$$(9) \qquad \vec{L} = \sum m_\alpha(\vec{r}_\alpha \times \vec{v}_\alpha)$$

as the <u>angular</u> <u>momentum</u> of the body (with respect to the fixed point); $\vec{r}_\alpha$ and $\vec{v}_\alpha$ are the position vector and the velocity vector respectively, of the α-th particle, m_α . Next, define the moment of inertia tensor as

$$(10) \qquad I = (I_{KL}) = ((\sum m_\alpha r_\alpha^2)\delta_{KL} - \sum m_\alpha x_{\alpha K} x_{\alpha L})$$

where the $x_{\alpha K}$ are either the components of $\vec{r}_\alpha$ in the space frame, in which case I need not be a constant in time, or in the body frame where I is necessarily constant. The expression for angular momentum (9) can now be written as

$$\vec{L} = I\vec{\omega} \quad \text{or} \quad L_K = I_{KL}\omega^L \quad . \tag{11}$$

The total <u>kinetic energy</u> of all the mass points is

$$T = \frac{1}{2}\sum m_\alpha v_\alpha^2 \quad , \tag{12}$$

or

$$T = \frac{1}{2}\vec{\omega}\cdot\vec{L} \quad , \tag{13}$$

or

$$T = \frac{1}{2} I_{KL}\,\omega^K\omega^L \quad . \tag{14}$$

We call T the <u>rotational kinetic energy</u>.

Suppose $\vec{x}_1$ and $\vec{x}_2$ denote two points of the body. The following formula relates the moment of inertia tensor $\underset{1}{I}$ at $\vec{x}_1$ to the corresponding $\underset{2}{I}$ at $\vec{x}_2$:

$$\left\{\begin{array}{l} \underset{1}{I}_{AB} = \underset{2}{I}_{AB} + m\,\eta^2\delta_{AB} - m\,\eta_A\eta_B \\[2ex] \vec{\eta} = \vec{x}_1 - \vec{x}_2 \end{array}\right. \tag{15}$$

(e) Euler's equations of motion for a rigid body held fixed at one point. The fundamental equation of motion for the total angular momentum

of a rigid body is

(16) $$\frac{d\vec{L}}{dt} = \vec{N}$$

where $\vec{N}$ is the total torque acting on the body, and all quantities necessarily referred to an inertial (space) frame, $(\vec{k}_\Gamma)$. In our point-mass model of a rigid body, $\vec{N}$ is usually defined to be

(17) $$\vec{N} = \sum \vec{r}_\alpha \times \vec{F}_\alpha \quad ,$$

where $\vec{F}_\alpha$ is the external force acting on the α-th particle (we do not include the forces of rigid constraint).

Via equation (8), we can transform (16) to its expression in the body frame:

(18) $$\left(\frac{d\vec{L}}{dt}\right)_{space} = \left(\frac{d\vec{L}}{dt}\right)_{body} + \vec{\omega} \times \vec{L}$$

which becomes, using the definition of $\vec{L}$ in (11)

(19) $$I\left(\frac{d\vec{\omega}}{dt}\right)_{body} + \vec{\omega} \times I\vec{\omega} = \vec{N}$$

(I is constant in the body frame!). These are Euler's equations. They take on a more familiar form if we write the components in the body frame $(\cdot \equiv \frac{d}{dt})$:

(20) $$I^A{}_B\, \dot{\omega}^B + \epsilon^A{}_{BC}\, \omega^B\, I^C{}_D\, \omega^D = N^A \quad ;$$

for example, if the principal axes of the moment of inertia tensor ([16], page 151) are chosen to be the body-fixed frame so that

$$I = \mathrm{diag}\,(I_1, I_2, I_3)$$

then Euler's equations are

$$I_1\,\dot{\omega}^1 + (I_2 - I_3)\omega^2\omega^3 = N^1$$

and the other two obtained by cyclically permuting the indices.

§2.3 Notations of Differential Geometry and Lie Groups.

(a) Manifolds. All manifolds will be assumed to be smooth (C^∞), paracompact, Hausdorff and connected. Such a manifold is often indicated by a capital Latin letter, M, N for example. The phrase "in local coordinates x^α" means that we have selected a chart $(U \subset M$, $x : U \to R^m)$ at some point $p \in M$ where M is m - dimensional and x^α $(1 \leq \alpha \leq m)$ is the α-th coordinate function of $x : U \to R^m$.

(b) Tangent bundle. If M is an m - dimensional manifold and $x \in M$ then T_xM denotes the tangent space to M at x; $TM = \bigcup_{x \in M} T_xM$ is then the tangent bundle of M with $\pi : TM \to M$, $(x,v) \to x$, the projection. T_x^*M and T^*M are the cotangent space at x, and cotangent bundle respectively.

Let N be a n - dimensional manifold and $\phi : M \to N$ a smooth map. We denote the differential of ϕ, a vector bundle map, by $\phi_* : TM \to TN$ and the dual of this map by $\phi^* : T^*N \to T^*M$. In a local coordinate system x^α of M and y^i of N, the coordinate basis of T_xM is $\{\partial_\alpha \equiv \frac{\partial}{\partial x^\alpha}\}$ and for T_x^*M is $\{dx^\alpha\}$ so that if $v = v^\alpha \frac{\partial}{\partial x^\alpha} \in$

T_xM then $\phi_*(v) = v^\alpha \frac{\partial\phi^i}{\partial x^\alpha}\frac{\partial}{\partial y^i}$ where ϕ^i are the components of ϕ and if $\sigma = \sigma_i\, dy^i \in T^*_{\phi(x)}N$, then $\phi^*(\sigma) = \frac{\partial\phi^i}{\partial x^\alpha}\sigma_i\, dx^\alpha$.

(1) Definition. A map $\phi : M \to N$ is a submersion at $x \in M$ iff $\phi_*|_x : T_xM \to T_xN$ has maximal rank of $\dim N \leq \dim M$

Let $y \in N$. If at each $x \in \phi^{-1}(y)$, ϕ is a submersion, then $\phi^{-1}(y)$ is an embedded submanifold of M ([59]).

(2) Definition. A map $\phi : M \to N$ is a diffeomorphism iff $\phi^{-1} : N \to M$ exists and is smooth.

(3) Definition. A map $\iota : M \to N$ is an embedding iff $\iota : M \to \iota(M) \subset N$ is a diffeomorphism.

(c) Vector fields and 1-forms. Recall that a vector field ξ is a smooth section of TM , $\xi : M \to TM$ such that $\pi\circ\xi = id_M$ and that a 1-form, σ , is a section of T^*M.[1)] Locally, $\xi(x) = \xi^\alpha(x) \frac{\partial}{\partial x^\alpha}\Big|_x$ and $\sigma(x) = \sigma_\alpha(x)\, dx^\alpha\Big|_x$ (often, we do not use $\Big|_x$ explicitly). We denote by $\mathfrak{X}(M)$ the set of all vector fields on M and by $\mathfrak{X}^*(M)$, the set of all 1-forms on M .

(d) Integral curves. Let $c : R \to M$ be a smooth curve. Denote by $\iota : R \to TR = R\times R$ the inclusion $t \to (t,1)$.

(4) Definition. The lift c' of c is a map $c' : R \to TM$ defined as $c' = c_*\circ\iota$. In local coordinates, if $t \to x^\alpha(t)$ is the curve in M , then c' is the curve $t \to (x^\alpha(t)\ ,\ \frac{dx^\alpha}{dt}\Big|_{x^\alpha(t)})$ in TM .

Let c be a smooth curve in M and ξ a vector field on M.

(5) Definition. c is an integral curve of ξ iff $c' = \xi \circ c$. Locally, this is just $\frac{dx^\alpha}{dt} = \xi^\alpha(x(t))$.

(e) Lie group. A Lie group G is a set with a group structure and a manifold structure compatible with the group operation (that is, the group operation is C^∞). If G is a group, $h, g \in G$ then $(h,g) \to hg$ denotes the group operation; e or id will denote the identity element of G.

The Lie algebra g of a Lie group G is the tangent space to G at the identity:

(6) $$g = T_e G \quad .$$

Let M be a manifold, G a Lie group.

(7) Definition. A map $\psi : G \times M \to M$ is called a (left) action of G on M iff $\psi(a) : M \to M$, $x \to \psi(a,x)$ is a diffeomorphism and $\psi(ab) = \psi(a) \circ \psi(b)$, $\psi(e) = id_M$.

It is evident that G can act on itself in three natural ways. Let $a \in G$ be fixed and b any element of G.

(8) Definition. Left translation by a in G is the map $\ell_a : G \to G$, $b \to ab$; similarly, right translation is the map $r_a : G \to G$, $b \to ba$. An inner automorphism of G by a is the map $i_a : G \to G$, $b \to aba^{-1}$.

Note the following properties of the induced differential maps:

(i) $(\ell_a)_*^{-1} = (\ell_{a^{-1}})_* : T_{ab}G \to T_bG$

so that

(ii) $(\ell_a)_* : T_bG \to T_{ab}G$ is an isomorphism.

The differential of the map $i_a : G \to G$, $(i_a)_* : T_bG \to T_{aba^{-1}}G$ is quite important. Consider its operation on g: clearly $(i_a)_* : g \to g$. This evidently gives us a representation of G in $\mathrm{Aut}(g)$, the set of all automorphisms of g. This representation is called the adjoint representation and denoted by

(9) $$a \to \mathrm{Ad}_a = (i_a)_*\big|_g \; .$$

Let $\xi \in \mathfrak{X}(G)$ be a vector field.

(10) Definition. ξ is said to be left-invariant iff for each $a \in G$, $(\ell_a)_*\xi = \xi \circ \ell_a$, i.e., ξ is ℓ_a - related to itself.

If ξ is a left-invariant vector field then

(11) $$(\ell_a)_*^{-1}(\xi\big|_a) = \xi\big|_e$$

so that this map {left-invariant vector fields} $\to T_eG = g$ is a Lie algebra isomorphism where the Lie bracket $[\xi,\eta] = \xi\eta - \eta\xi$ defines the Lie algebra structure on the left-invariant vector fields.[1)]

Let $\sigma \in \mathfrak{X}^*(G)$.

(12) Definition. σ is a left-invariant 1-form (or Maurer-Cartan form) iff for all $a \in G$, $(\ell_a)^* \sigma = \sigma$.

Let $\{X_1, \ldots, X_n\}$ be a basis of g where $n = \dim G$; $\{X^1, \ldots, X^n\}$ denotes the dual basis for $g^* = T_e^* G$. There exist constants C^K_{AB} such that

(13)
$$[X_A, X_B] = C^K_{AB} X_K$$

and are called the structure constants of G with respect to the basis $\{X_K\}$ of g.

If $A \in g$, denote the corresponding left-invariant vector field by $\underset{\sim}{A} \in \mathfrak{X}(G)$; similarly, if $\sigma \in g^*$, denote the corresponding left-invariant form by $\tilde{\sigma}$. It then follows that TG is parallelizable by $\{\underset{\sim}{X}_K\}$: $TG \cong G \times g$; similarly $T^*G \cong G \times g^*$, The Maurer-Cartan equations are

(14)
$$d\tilde{X}^A = \frac{1}{2} C^A_{BC} \tilde{X}^B \wedge \tilde{X}^C .$$

(f) Exponential map.

(15) Definition. A homomorphism $\phi : R \to G$ is called a 1-parameter subgroup of G.

Let $\lambda \to \exp_A(\lambda)$ be the unique 1-parameter subgroup of G such that $\left.\frac{d}{d\lambda}\right|_{\lambda=0} (\exp_A(\lambda)) = A$ where $A \in g$.

(16) Definition. The exponential map $\exp : g \to G$ is defined by

$\exp(A) = \exp_A(1)$. Further details and properties may be found in [59].

In the case that G is a matrix group, $\exp$ is given by matrix exponentiation: if $A \varepsilon g$, then

$$\exp(A) = e + A + \frac{1}{2!} A^2 + \cdots$$

where $e \varepsilon G$ again denotes the identity element.

(g) Example: the general linear group. $GL(n) \equiv GL(n,R)$ is the group of all invertible linear transformations of $R^n \to R^n$. Let $(x^\alpha ; \alpha = 1,\ldots,n)$ be the standard coordinate system on R^n . We identify each $a \varepsilon GL(n)$ with its matrix representative relative to the basis $\{\partial_\alpha\}$ of R^n : let $X = X^\alpha \partial_\alpha \varepsilon R^n$; then $a(X) = X^\alpha a(\partial_\alpha) = X^\alpha a^\beta{}_\alpha \partial_\beta$. According to this last expression, we can view a either as mapping the basis $\{\partial_\alpha\}$ to the basis $\{a^\beta{}_\alpha \partial_\beta\}$ keeping the components with respect to each basis the same, or reciprocally, the basis $\{\partial_\alpha\}$ remains fixed and the components X^α change to $a^\alpha{}_\beta X^\beta$. We examine these actions in more detail in the Appendix.

Left translation in $GL(n)$ is accomplished by matrix multiplication: let $a \varepsilon GL(n)$; then $\ell_a(b) = a \cdot b$ for any $b \varepsilon GL(n)$. In components, $(\ell_a(b))^\alpha{}_\beta = a^\alpha{}_\rho b^\rho{}_\beta$. Similarly, let $v \varepsilon T_b\, GL(n)$; then $(\ell_a)_*(v) = av \varepsilon T_{ab}\, GL(n)$.

The Lie algebra of $GL(n)$, $g\ell(n) = T_e GL(n)$ is diffeomorphic to $R^{n\times n}$, the space of all $n\times n$ matrices. The result of multiplication in $g\ell(n)$ of two elements A and B is their commutator:

$[A,B] = AB-BA \in g\ell(n)$. We choose as standard basis for $g\ell(n)$ the matrices $\{E^i_j\}$ defined by

$$(17) \qquad (E^i_j)^a{}_b = \delta^i_b \, \delta^a_j \quad .$$

Denote by (e^α_i) the standard coordinate functions on $GL(n)$ and by (θ^i_α) the inverse to (e^α_i) . Then,

$$(18) \qquad \underset{\sim}{E}^i_j = e^\alpha_j \frac{\partial}{\partial e^\alpha_i} \quad \text{and}$$

$$(19) \qquad \tilde{E}^i_j = \theta^i_\alpha \, de^\alpha_j \quad .$$

§2.4 The Rotation Group in Three Dimensions.

(a) In §2.2 it was argued that the group $SO(3)$ serves to describe the configuration manifold of a rigid body held fixed at one point. In this section we examine in more detail $SO(3)$ and its role in the description of a rigid body.

Capital Greek and Latin indices both run and sum over 1, 2, and 3. We assume that a global space frame $\{\ell_r\}$ is defined at each point of R^3 in accordance with Euclidean parallelism; $\{\eta^r\}$ denotes the dual basis.

We recall from §2.1 that

$$(1) \qquad SO(3) = \{a \in GL(3) : aa^t = id \ , \ \det a = 1\} \quad .$$

$SO(3)$ is a closed Lie subgroup of $GL(3)$ and we consider it as a

subset of GL(3) ; in particular, the usual coordinate functions e^{Γ}_{A}, the coordinate vector fields $\frac{\partial}{\partial e^{\Gamma}_{A}}$ and 1-forms de^{Γ}_{A} will be used as for GL(3) .[2)]

(b) Angular velocity. Let

(2) $$c : R \to SO(3) \quad , \quad t \to e^{\Gamma}_{A}(t)$$

be a smooth curve in SO(3) , representing a motion of the body. The tangent vector to the curve, $\dot{e}(t) = \dot{e}^{\Gamma}_{A}(t) \frac{\partial}{\partial e^{\Gamma}_{A}}$ $(\cdot \equiv \frac{d}{dt})$, can be written as

(3) $$\dot{e}(t) = \dot{e}^{\Gamma}_{A}\, \theta^{B}_{\Gamma}\, \underset{\sim}{E}^{A}_{B} \quad ,$$

showing that the quantity

(4) $$\begin{cases} \Omega \equiv \Omega_{A}{}^{B}\, E^{A}_{B} \\ \\ \Omega_{A}{}^{B} \equiv \dot{e}^{\Gamma}_{A}\, \theta^{B}_{\Gamma} \end{cases}$$

is an element of the Lie algebra so(3) of SO(3) and that

(5) $$\dot{e}(t) = \underset{\sim}{\Omega} \; .$$

To see the physical meaning of Ω , recall that the body and space frames are related by

(6) $$e_{A} = e^{\Gamma}_{A}\, \ell_{\Gamma}$$

which defines the orientation of the body at time t. Then

$$\dot{e}_A = \Omega_A{}^B e_B \ , \tag{7}$$

defining the (1,1) tensor

$$\bar{\Omega} = \Omega_A{}^B e_B \otimes \theta^A \ . \tag{8}$$

Transforming to the space frame, we see that

$$\bar{\Omega} = (\theta^A_\Gamma \Omega_A{}^B e^\Lambda_B) \ell_\Lambda \otimes \eta^\Gamma \ , \tag{9}$$

defining the space components of $\bar{\Omega}$:

$$\Omega_\Gamma{}^\Lambda = \theta^A_\Gamma \dot{e}^\Lambda_A \ . \tag{10}$$

Equation (7) is the statement that $\bar{\Omega}$ is the angular velocity tensor; (8) is its expression in the body frame while (9) is its expression in the space frame. The associated angular velocity vector 2.2 (7) is

$$\left\{\begin{array}{ll} \hat{\Omega} = \Omega^A e_A = -\frac{1}{2}\epsilon^{AB}{}_C \Omega_B{}^C e_A & \text{(body frame)} \\ \\ \hat{\Omega} = \Omega^\Gamma \ell_\Gamma = -\frac{1}{2}\epsilon^{\Gamma\Lambda}{}_\Delta \Omega_\Lambda{}^\Delta \ell_\Gamma & \text{(space frame)} \end{array}\right. \tag{11}$$

To recapitulate: we have defined two tensors, Ω from (3) and (4) and $\bar{\Omega}$ from (7) and (8) which lie in completely different spaces: $\Omega \in so(3)$ while $\bar{\Omega} \in T^1_1(R^3)$. $\bar{\Omega}$ is the usual angular velocity tensor and has exactly the same components as Ω; whence, in our model

of rotational motion (being $SO(3)$) Ω is the angular velocity tensor in the body frame.

(c) Lie algebra of $SO(3)$. By differentiating the defining relations (1), one obtains

(12) $$so(3) = \{A \in g\ell(3) : A + A^t = 0\} .$$

A suitable basis for $so(3)$ is then $\{E_{AB} \equiv E^A_B - E^B_A ,\ 1 \leq A < B \leq 3\}$; more often, we shall employ the matrices

(13) $$E_A = -\frac{1}{2} \epsilon_A{}^{BC} E_{BC} = -\epsilon_{AB}{}^{C} E^B_C .$$

Let $\{E^A\}$ be the dual basis of g^*.

(d) We now give a more formal description of (c) in $SO(3)$. Recalling the definitions in §2.3 and especially Example (g), equation (4) is

(14) $$\Omega = (\ell_{e^{-1}})_* (\dot{e}) \in g$$

and if we identify Ω with $\overline{\Omega}$ then

(15) $$\begin{cases} \Omega_{body} = Ad_e (\Omega_{space}) \\ \\ \Omega_{body} = (\Omega_A{}^B) \ ; \ \Omega_{space} = (\Omega_\Lambda{}^\Gamma) . \end{cases}$$

(e) Coordinates. Recall (§2.3e) that $TSO(3) \cong SO(3) \times so(3)$. The left-invariant vector fields $\{\underset{\sim}{E}_A\}$ are a basis of $T_a SO(3)$ for each $a \in SO(3)$, so that (e^Γ_A, Ω^A) forms a coordinate system for $TSO(3)$. We also have the coordinate expressions

$$(16) \qquad \underset{\sim}{E}_A = -\epsilon_{AB}{}^C \, e^\Gamma_C \frac{\partial}{\partial e^\Gamma_B}$$

$$(17) \qquad \left\{ \begin{array}{l} \tilde{E}^A = -\epsilon^{AK}{}_L \, \theta^L_\Gamma \, de^\Gamma_K \\ \text{iff} \\ de_A = -e^\Gamma_K \, \epsilon^K{}_{AL} \, \tilde{E}^L \end{array} \right.$$

(f) Maurer-Cartan equations. Relative to the basis $\{E_A\}$ of $so(3)$ the structure constants of $SO(3)$ are $C^A{}_{BC} = \epsilon^A{}_{BC}$ so that the Maurer-Cartan equations are

$$(18) \qquad d\tilde{E}^A = -\frac{1}{2}\epsilon^A{}_{BC} \, \tilde{E}^B \wedge \tilde{E}^C \; .$$

(g) Left-invariant metrics. A left-invariant positive definite metric I on a Lie group G is a $(2,0)$ - tensor field such that I_a is a positive definite symmetric bilinear form on $T_a G$ for all $a \in G$ and satisfies $(\ell_a)^* I = I$. Note that this implies that I is obtainable from a positive definite symmetric bilinear form I_{id} on g: $I_a = (\ell_{a^{-1}})^* I_{id}$. Denote I_{id} by $\tilde{I}$ and consider $\tilde{I}$ as an isomorphism $g \to g^*$ so that

$$(19) \qquad I_a(v) = (\ell_{a^{-1}})^* \tilde{I}((\ell_{a^{-1}})_*(v)) \in T^*_a G \quad \text{if} \quad v \in T_a G \; .$$

The metric I determines a function $L : TG \to R$ defined by

$$L(a,v) = \frac{1}{2} I_a(v,v) \quad . \tag{20}$$

The following diagram is a useful summary:

$$\begin{tikzcd}
\mathfrak{g} \arrow[rr, "Ad_a"] \arrow[ddd, "\tilde{I}"'] & & \mathfrak{g} \\
& T_aG \arrow[ul, "(\ell_{a^{-1}})_*"'] \arrow[ur, dashed, "(r_{a^{-1}})_*"] \arrow[d, "I_a"] & \\
& T^*_aG \arrow[dl, dashed, "\ell^*_a"'] \arrow[dr, dashed, "(r_a)^*"] & \\
\mathfrak{g}^* & & \mathfrak{g}^* \arrow[ll, "Ad^*_a"]
\end{tikzcd} \tag{21}$$

In our $SO(3)$ model of the rigid body, I is the moment of inertia tensor (necessarily with respect to the body frame), and has the form

$$I = I_{AB}\, E^A \otimes E^B \quad . \tag{22}$$

Now, $\Omega = \Omega^A E_A$ is the angular velocity; the quantity

$$S = \tilde{I}(\Omega) = S_A E^A = I_{AB}\, \Omega^B E^A \tag{23}$$

or (amounting to the same thing)

$$\underset{\sim}{S} = I(\underset{\sim}{\Omega}) = S_A \underset{\sim}{E}^A \tag{24}$$

is the spin angular momentum of the rigid body. This is justified by writing 2.2 (11) in terms of the body frame (e_A) : $L = L_A \theta^A = I_{AB}\Omega^B \theta^A$ and noticing that the components $L_A = S_A$.

The function defined by (19) is now written

$$L(a,\Omega) = \frac{1}{2} I_a(\Omega,\Omega) = \frac{1}{2} I_{AB} \Omega^A \Omega^B \tag{25}$$

with the identification $\tilde{I} \leftrightarrow I$ and $\underset{\sim}{\Omega} \leftrightarrow \Omega$ if there is no confusion. We recognize (25) as the usual expression 2.2 (14) for the rotational kinetic energy.

§2.5 Canonical Lagrangian Dynamics.

(a) The modern geometrical formulation of canonical Lagrangian and Hamiltonian mechanics can be found in several books devoted to the subject (Abraham [1], Souriau [48]; Chernoff and Marsden [7] for infinite-dimensional systems) as well as books on differential geometry ([15], [51], for example). This section derives a Lagrangian formulation from the classical Euler-Lagrange equations, then sketches the general inhomogeneous formalism.

(b) Let Q be the configuration space of some physical system with n degrees of freedom (so that $\dim Q = n$). TQ is then the state space (or velocity phase space) of the system. Locally, we have coordinates (q^α) on Q and (q^α, v^α) on TQ . Let $L : TQ \to R$ be a Lagrangian describing the dynamics of the system. Classically, a motion of the system is represented by a curve $t \to (x^\alpha(t), v^\alpha(t))$ in state space which satisfies

(1) $$v^{\alpha} = \frac{dx^{\alpha}}{dt}$$

and the Euler-Lagrange equations

(2) $$\frac{d}{dt}\left(\frac{\partial L}{\partial v^{\alpha}}\right) - \frac{\partial L}{\partial q^{\alpha}} = 0 \quad .$$

We first postulate that $t \to (x^{\alpha}(t), v^{\alpha}(t))$ is an integral curve of a vector field $X \in \mathfrak{X}(TQ)$. Then,

(3) $$v^{\alpha} = \frac{dq^{\alpha}}{dt} = X \lrcorner\, dq^{\alpha} = X(q^{\alpha}) \quad ,$$

(4) $$\frac{d}{dt}\left(\frac{\partial L}{\partial v^{\alpha}}\right) = X\left(\frac{\partial L}{\partial v^{\alpha}}\right) \quad ,$$

whence (2) is equivalent to

(5) $$X\left(\frac{\partial L}{\partial v^{\alpha}}\right) dq^{\alpha} - \frac{\partial L}{\partial q^{\alpha}}\, dq^{\alpha} = 0 \quad .$$

Expressing everything in terms of Lie derivation by X ,

$$L_X\left(\frac{\partial L}{\partial v^{\alpha}}\right) = X\left(\frac{\partial L}{\partial v^{\alpha}}\right) \quad ,$$

$$L_X\left(\frac{\partial L}{\partial v^{\alpha}}\, dq^{\alpha}\right) = L_X\left(\frac{\partial L}{\partial v^{\alpha}}\right) dq^{\alpha} + \frac{\partial L}{\partial v^{\alpha}}\, L_X(dq^{\alpha}) \quad \text{and}$$

$$L_X(dq^{\alpha}) = d(X \lrcorner\, dq^{\alpha}) = dv^{\alpha}$$

and defining the 1-form

(6) $$\theta_L = \frac{\partial L}{\partial v^{\alpha}}\, dq^{\alpha} \quad ,$$

equation (5) becomes:

(7) $$L_X(\theta_L) - [\frac{\partial L}{\partial q^\alpha} dq^\alpha + \frac{\partial L}{\partial v^\alpha} dv^\alpha] = 0 \quad .$$

We define two more quantities:

(8) $$\omega_L = - d\theta_L \quad ,$$

(9) $$E = \frac{\partial L}{\partial v^\alpha} v^\alpha - L \quad .$$

Since $L_X(\theta_L) = d(X \lrcorner \theta_L) + X \lrcorner d\theta_L = d(\frac{\partial L}{\partial v^\alpha} v^\alpha) - X \lrcorner \omega_L$ then we obtain the final form:

(10) $$X \lrcorner \omega_L - dE = 0 \quad .$$

If L is non-degenerate (that is $\frac{\partial^2 L}{\partial v^\alpha \partial v^\beta}$ is nonsingular on TQ) then ω_L is a sympletic form.[3)] E is the energy of the system.

The steps leading to (10) are reversible, showing that the classical formulation (1), (2) is equivalent to the symplectic formulation (10).

(c) Canonical mechanics. Again Q denotes configuration space, TQ state space, T^*Q (momentum) phase space and $L : TQ \to R$ a Lagrangian. Let (q^α, p_α) be a local coordinate system on T^*Q .

(11) Theorem. ([51] pg. 143). There is a 1-form $\theta_o \in \mathcal{X}^*(T^*Q)$ defined by the property: for each $u \in T^*Q$ and for each $X \in T_u T^*Q$ we have $\langle X, \theta_o|_u \rangle = \langle \pi_* X, u \rangle$ where $\pi : T^*Q \to Q$ is the bundle projection and $\langle , \rangle$ is the natural pairing between $T_u T^*Q$ and $T_u^* T^*Q$ in one case and TQ and T^*Q in the second. The 2-form $\omega_o = - d\theta_o$ has maximal rank and

thus defines a symplectic structure on T^*Q . In coordinates (q^α, p_α) on T^*Q , $\theta_o = p_\alpha dq^\alpha$ and $\omega_o = dq^\alpha \wedge dp_\alpha$.

Let $L \in C^\infty(TQ)$ and L_q denote $L|_{T_qQ}$.

(12) __Definition__. The __fibre derivative__ of L is a map $FL : TQ \to T^*Q$ defined by: for each $u = (q,v) \in TQ$, $FL(u) = (q,p)$ where $p \in T_q^*Q$ has the following expression:

$$p(w) = \frac{d}{ds}\Big|_{s=0} L_q(v+sw) \quad \text{for all} \quad w \in T_qQ \quad .$$

In a local coordinate system (q^α) , $v = v^\alpha \dfrac{\partial}{\partial q^\alpha}$ and $FL(v) = \dfrac{\partial L}{\partial v^\alpha} dq^\alpha$.

(13) __Definition__. L is a __regular Lagrangian__ iff $FL : TQ \to T^*Q$ is a regular map on TQ (__ie__, dFL is nowhere zero on TQ).

(14) __Proposition__ ([1], pg. 116). L is a regular Lagrangian iff FL is a local diffeomorphism iff $FL^*(\omega_o)$ is a symplectic form on TQ .

We remark that, in general, the fibre derivative FL is not linear as a map on the fibres of TQ . If $v \in T_qQ$ is the velocity, then $p = FL(v)$ is the __canonical__ (or __conjugate__) __momentum__ associated to v . The symplectic manifold (T^*Q, ω_o) is the arena of Hamiltonian mechanics while $(TQ, \omega_L = FL^*(\omega_o))$ is that of canonical Lagrangian dynamics.

We need two further definitions:

(15) __Definitions__. Let L be a regular Lagrangian. The function $A : TQ \to R$, $A(q,v) = \langle v, FL(v) \rangle$ is the __action__ of the Lagrangian

L . The function $E = A-L$ is the energy of L . In coordinates (q^α, v^α) on TQ ,

(16)
$$A = \frac{\partial L}{\partial v^\alpha} v^\alpha$$

(17)
$$E = \frac{\partial L}{\partial v^\alpha} v^\alpha - L \quad .$$

Let L be a regular Lagrangian. The symplectic form $\omega_L = FL^*(\omega_o)$ on TQ has maximal rank and thus provides a linear isomorphism between $T_u TQ$ and $T_u^* TQ$ for each $u \in TQ$. Denote by $\tilde{\omega}$ the induced isomorphism between $\mathfrak{X}(TQ)$ and $\mathfrak{X}^*(TQ)$ so that if X is any vector field on TQ then $\tilde{\omega}(X)$ is a 1-form on Q . The canonical Euler-Lagrange equations are

(18)
$$X = \tilde{\omega}^{-1}(dE)$$

or equivalently,

(19)
$$X \lrcorner \omega = dE$$

which is seen to be the same as in (10). The motions of the system are represented by integral curves in TQ of the vector field X satisfying (19). It is shown in ([1], pg. 120) that X defined by (18) is a second-order equation on Q :

(20) Definition. A second-order equation on a manifold M is a vector field X on TM such that $\pi_* \circ X = id_{TM}$ where $\pi : TM \to M$ and $\pi_* : TTM \to TM$ is its differential. In a local coordinate system (x^α, v^α) of TM , X must have the form

$$X = v^\alpha \frac{\partial}{\partial x^\alpha} + f^\alpha(x,v) \frac{\partial}{\partial v^\alpha} \quad . \tag{21}$$

Last, we remark that the Lagrangian structure on TQ induced by the canonical structure on T^*Q is not the only way a Lagrangian formalism can be formulated. An alternative approach, making use of the so-called almost tangent structure on TQ, is used by [23] and [58] and discussed briefly in §5.3.

§2.6 Canonical Dynamics of Rotational Motion of a Rigid Body.

(a) The canonical Lagrangian formalism described in the last section is now applied to the configuration space $SO(3)$ representing rotational motion of a rigid body as formulated in §2.4. Henceforth:

$$Q = SO(3)$$

$$TQ = SO(3) \times so(3)$$

(1)

$$L : TQ \to R \ , \ L(q,\Omega) = \frac{1}{2} I_q(\underset{\sim}{\Omega},\underset{\sim}{\Omega}) = \frac{1}{2} I_{AB}\Omega^A\Omega^B$$

$$FL : TSO(3) \to T^*SO(3) \ , \ FL(\Omega)_A = S_A \equiv I_{AB}\Omega^B$$

$$\theta_o \in \mathcal{X}^*(T^*SO(3)) \ , \ \theta_o = S_A\tilde{E}^A$$

$$\theta_L = FL^*(\theta_o) = I_{AB}\Omega^B \tilde{E}^A$$

The canonical symplectic form $\omega_L = -d\theta_L$ can be computed using the Maurer-Cartan equations 2.4(18):

$$\omega_L = -I_{AB}\, d\Omega^A \wedge \tilde{E}^B + \frac{1}{2} S_{AB}\, \tilde{E}^A \wedge \tilde{E}^B \tag{2}$$

where we still write $S_A = I_{AB}\,\Omega^B$ for convenience and

$$S_{AB} = \epsilon_{AB}{}^{C}\, S_C \quad . \tag{3}$$

The energy (see 2.5 (17)) is

$$E = \frac{1}{2}\, I_{AB}\, \Omega^A \Omega^B \tag{4}$$

so that

$$dE = S_A\, d\Omega^A \quad . \tag{5}$$

Let $X \in \mathfrak{X}(TSO(3))$ be any vector field, say,

$$X = P^A \underset{\sim}{E}_A + \xi^A \frac{\partial}{\partial \Omega^A} \quad . \tag{6}$$

Then

$$X \lrcorner\, \omega_L = -(S_{AB}P^B + I_{AB}\xi^B)\tilde{E}^A + I_{AB}P^B\, d\Omega^A \tag{7}$$

so that equation 2.5 (19), $X \lrcorner\, \omega = dE$, implies

$$P^A = \Omega^A \quad \text{and} \tag{8}$$

$$\left\{ \begin{array}{l} \xi^A = -\tilde{I}^{AB} S_{BC} \Omega^C \\ \\ (\tilde{I}^{AB}) \text{ is the inverse to } (I_{AB}) \quad . \end{array} \right. \tag{9}$$

Note that X given by (6), (8) and (9) is of the form 2.5 (20), a second-order equation.

(c) Integral curves. Let $t \to (e^{\Gamma}_{A}(t), \Omega^{A}(t))$ be an integral curve of X ; physically, this integral curve represents the evolution of the system with respect to time t . Then,

$$\frac{de^{\Gamma}_{A}}{dt} = X \lrcorner\, de^{\Gamma}_{A} = X \lrcorner (-e^{\Gamma}_{K}\, \epsilon^{K}{}_{AL}\, \tilde{E}^{L}) = -\epsilon_{AB}{}^{C}\, \Omega^{B}\, e^{\Gamma}_{C} \tag{10}$$

where we have used 2.4 (17). Equation (10) is merely a regurgitation of the definition 2.4 (7).

Last, for angular velocity we have

$$\frac{d\Omega^{A}}{dt} = X \lrcorner\, d\Omega^{A} = -\tilde{I}^{AB}\, S_{BC}\, \Omega^{C} \tag{11}$$

which we recognize as Euler's equations, 2.2 (20).

CHAPTER III

A Canonical Lagrangian Formalism for a Nonrelativistic Rigid Body

§3.1 Introduction.

We saw in Chapter I how the motion of a rigid body held fixed at one point can be modeled as a canonical Lagrangian system on the Lie group $SO(3)$. The generalization to include motion in space according to a fairly general Lagrangian is made in this Chapter, partly for its intrinsic interest, partly to develop more formalism as preparation for the space-time formalisms which will be met in Chapters IV and V.

In simple situations (for example, a gyroscope falling in a gravitational field) the rotational and translational motions occur independently and one takes $R^3 \times SO(3)$ as configuration space with a Lagrangian such as: $L : T(R^3 \times SO(3)) \to R$, $(x,e;v,\Omega) \to \frac{1}{2} mv^2 + \frac{1}{2} I_{AB}\Omega^A\Omega^B - m\phi(x)$, m being the mass of the body. However, we shall generalize in two directions: First, space shall be an arbitrary 3-dimensional Riemannian manifold Σ , representing the configuration manifold of a point of reference fixed in the body (usually the centre of mass). The physical significance of allowing Σ to be so general is not clear; however, it may have some interpretation in the mechanics of generalized continua. Our main purpose though, in the generality of Σ , is having the present formalism as a point for comparison with the later space-time theories. The second generalization has already been mentioned: a Lagrangian which can include electromagnetic-type interactions and spin-orbit like interactions as special cases.

§3.2 Bundle of Orthonormal Triads as Configuration Space.

(a) It was remarked earlier (§2.1) how the specification of a body fixed triad of vectors (with respect to a set of space axes having the same origin - more about this later) leads to the description of the rotational motion of a rigid body by $SO(3)$. Thus our representation of the general configuration of the body consists of specifying

(i) a point $x \in \Sigma$, a 3-dimensional Riemannian manifold with metric tensor γ : (centre of mass).

(ii) an orthonormal frame $e = (e_A)$ of $T_x \Sigma$: (orientation with respect to centre of mass).

The totality of all orthonormal frames at all points in Σ leads to the choice of $SO(\Sigma)$, a principal fibre bundle over Σ with $SO(3)$ as structure group, as configuration space for the rigid body (see the Appendix for the definitions, or [24] for a complete discussion of principal bundles).

(b) $SO(\Sigma)$: definitions. $SO(\Sigma)$ is a sub-bundle of $L(\Sigma)$ and is characterized by the metric on Σ : if $(x,e) \in L(\Sigma)$, then

(1) $$(x,e) \in SO(\Sigma) \quad \text{iff} \quad \gamma_x(e_A, e_B) = \delta_{AB} \ .$$

It is clear that if $a \in SO(3)$ then

(2) $$e \in SO_x\Sigma \quad \text{iff} \quad ea \in SO_x\Sigma \ ,$$

and this shows that $SO_x\Sigma \cong SO(3)$.

This process of obtaining $SO(\Sigma)$ from $L(\Sigma)$ is called <u>reducing</u> $L(\Sigma)$ <u>to</u> $SO(\Sigma)$; $SO(\Sigma)$ is one of several examples of G - structures that we will be meeting (cf: [51] for the general theory).

A local coordinate system on a neighbourhood U of $u = (x,e) \in SO(\Sigma)$ will still be denoted by (x^Γ, e^Γ_A) (see the Appendix) but with $e = (e_A)$ subject to the results of §2.4.

The linear connection Γ on Σ that we will use is the ordinary Levi-Civita connection of the metric γ:

(3) $$\Gamma^\Gamma_{\Lambda\Delta} \equiv \{^{\Gamma}_{\Lambda\Delta}\} \quad .$$

The $so(3)$ - valued connection 1-form is

(4) $$\left\{ \begin{aligned} \omega &= \tfrac{1}{2}\omega^{AB} \, E_{AB} \\ \omega^A{}_B &= \theta^A(de^\Gamma_B + \Gamma^\Gamma_{\Lambda\Delta} \, e^\Delta_B \, dx^\Lambda) \\ E_{AB} &= E^A_B - E^B_A \, , \quad A < B. \end{aligned} \right.$$

The canonical 1-form is essentially the same as A(7):

(5) $$\left\{ \begin{aligned} \theta &= (\theta^A) \\ \theta^A &= \theta^A_\Gamma \, dx^\Gamma \quad , \quad \theta^A_\Gamma = (e^{-1})^A_\Gamma \end{aligned} \right.$$

at a point $(x,e) \in SO(\Sigma)$.

We take $\{\underset{\sim}{e}_A, \underset{\sim}{E}_A\}$ as a basis for $\mathfrak{X}(SO(\Sigma))$ with

(6) $$\underset{\sim}{e}_A = e^\Gamma_A \frac{\partial}{\partial x^\Gamma} - \Gamma^\Gamma_{\Lambda\Delta}\, e^\Delta_B\, e^\Lambda_A \frac{\partial}{\partial e^\Gamma_B} \quad \text{and}$$

(7) $$\underset{\sim}{E}_A = -\frac{1}{2}\epsilon_A{}^{BC}\,\underset{\sim}{E}_{BC} = -\epsilon_{AB}{}^C\, e^\Gamma_C \frac{\partial}{\partial e^\Gamma_B}$$

and the corresponding dual basis is denoted by $\{\theta^A, \tilde{E}^A\}$.

(c) The interpretation of the various quantities is essentially the same as in §2.4. Ω^A are the components of the angular velocity in the body frame e at a point $(x,e) \in SO(\Sigma)$ when the centre of mass is at x while $v^A \underset{\sim}{e}_A$ projects (under $\pi_* : TSO(\Sigma) \to T\Sigma$) to the velocity of the centre of mass: $\pi_*(v^A \underset{\sim}{e}_A) = v^A e^\Gamma_A \frac{\partial}{\partial x^\Gamma} \in T_x\Sigma$.

Orientation of the body has been defined with respect to the coordinate basis $\{\frac{\partial}{\partial x^\alpha}\}$ (at least locally). In a general setting it is difficult to determine a priori whether the body is rotating as the centre of mass moves through space since $\Gamma \neq 0$. After the path c in Σ of the centre of mass has been determined, one can compare the behavior of the motion $t \to e(t)$ to a chosen frame field which is parallely propagated along c , at least in the present situation.

(d) We list for reference the structure equations of the connection (see A(15) and A(16)) in this case:

(8) $$d\theta^A = -\epsilon^A{}_{BC}\, \tilde{E}^B \wedge \theta^C$$

(9) $$d\tilde{E}^A = -\frac{1}{2}\epsilon^A{}_{BC}\left(\tilde{E}^B \wedge \tilde{E}^C + \frac{1}{2} R^{BC}{}_{KL}\,\theta^K \wedge \theta^L\right)$$

where $\frac{1}{2} R^{BC}{}_{KL}\, \theta^K \wedge \theta^L = \Omega^{BC}$ is the $so(3)$ - valued curvature 2-form of the connection.

§3.3 Canonical Dynamics of the General Motion.

(a) The programme of canonical Lagrangian mechanics outlined in §2.5 is now applied to the configuration space $Q = SO(\Sigma)$. The Riemannian metric on Σ is written as either $\gamma = \gamma_{\Lambda\Gamma}\, dx^{\Lambda} \otimes dx^{\Gamma}$ or $\gamma = \delta_{AB}\, \theta^{A} \otimes \theta^{B}$. We assume the rotational motion is described as in §2.6.

The Lagrangian $L : TSO(\Sigma) \to R$ is

(1) $$L = \frac{1}{2} m\, \delta_{AB} v^{A} v^{B} + \frac{1}{2} I_{AB} \Omega^{A} \Omega^{B} + K_{AB} v^{A} \Omega^{B} + A_{A} v^{A} + B_{A} \Omega^{A} - m\phi$$

where I_{AB} is the moment of inertia tensor with respect to the centre of mass,

(2) $$T_{trans} = \frac{1}{2} m\, \delta_{AB} v^{A} v^{B} \quad : \quad \text{translational kinetic energy}$$

(3) $$T_{rot} = \frac{1}{2} I_{AB} \Omega^{A} \Omega^{B} \quad : \quad \text{rotational kinetic energy}$$

and (2) and (3) are both defined with respect to the centre of mass. The term $K_{AB} v^{A} \Omega^{B}$, with K_{AB} depending on x only, represents a sort of "spin-orbit" coupling while $A_{A} v^{A}$ and $B_{A} \Omega^{A}$ are velocity-dependent potentials (A_{A}, and B_{A} are functions of x only). Finally, $m\phi$ is the potential energy and again ϕ is a function of x only.

(b) Special cases. Unfortunately, the Lagrangian (1) is too general to generate the corresponding equations of motion in a reasonable amount of space. To illustrate the method and typical results, we exam-

ine two special cases:

Case (i):

(4) $$L = \frac{1}{2} m \, \delta_{AB} v^A v^B + \frac{1}{2} I_{AB} \Omega^A \Omega^B + K_{AB} v^A \Omega^B$$

where K is constant, and is "small" compared to $m^{-1} I_{AB}$.

Case (ii):

(5) $$L = \frac{1}{2} m \, \delta_{AB} v^A v^B + \frac{1}{2} I_{AB} \Omega^A \Omega^B + m \, A_A v^A + m \, B_A \Omega^A - m\phi$$

where A , B and ϕ are functions depending on $x \in \Sigma$ only.

(c) Canonical 1-form for the Lagrangian (1). Define the quantities

(6) $$\Theta_A = mv_A + K_{AB}\Omega^B + A_A \quad ,$$

(7) $$\textstyle\sum_A = I_{AB}\Omega^B + K_{BA} v^B + B_A \quad ;$$

these are the momenta conjugate to v_A and Ω_A respectively. Then,

(8) $$\theta = \Theta_A \theta^A + \textstyle\sum_A \tilde{E}^A$$

is the canonical 1-form of the Lagrangian (1).

(d) Case (i). We first compute the energy E associated to the Lagrangian L given by (4):

(9) $$E = v^A \Theta_A + \Omega^A \textstyle\sum_A - L$$

$$= \frac{1}{2} m\delta_{AB} v^A v^B + \frac{1}{2} I_{AB}\Omega^A\Omega^B + K_{AB}\Omega^B v^A \quad ,$$

so that

(10) $$dE = (mv_A + K_{AB}\Omega^B)\, dv^A + (I_{AB}\Omega^B + K_{BA}v^B)\, d\Omega^A \quad .$$

From the 1-form (8) we deduce the symplectic form $\omega = -d\theta_L$,

(11) $$\omega = -m\, dv_A \wedge \theta^A + K_{AB}\theta^A \wedge d\Omega^B + \epsilon_{ABC}\Theta^A\, \tilde{E}^B \wedge \theta^C$$

$$- I_{AB}\, d\Omega^A \wedge \tilde{E}^B + K_{BA}\, \tilde{E}^A \wedge dv^B$$

$$+ \frac{1}{2} \epsilon_{ABC} \Sigma^A\, \tilde{E}^B \wedge \tilde{E}^C + \frac{1}{4} \epsilon_{ABC} \Sigma^A R^{BC}{}_{KL}\theta^K \wedge \theta^L \quad .$$

Let

(12) $$Z = \nu^A e_A + P^A E_A + \xi^A \frac{\partial}{\partial v^A} + \Phi^A \frac{\partial}{\partial \Omega^A}$$

be a typical vector field on $TSO(\Sigma)$. Using (11) and (12), the 1-form $Z \lrcorner\, \omega$ is found to be

(13) $$Z \lrcorner\, \omega = -(K_{AB}\Phi^B + m\xi_A - \epsilon_{ABC}\Theta^B P^C + \frac{1}{2} \epsilon_{LMN}\Sigma^L R^{MN}{}_{AB}\nu^B)\theta^A$$

$$-(I_{AB}\Phi^B + K_{BA}\xi^B - \epsilon_{ABC}\Sigma^B P^C - \epsilon_{ABC}\Theta^B \nu^C)\tilde{E}^A$$

$$+(m\nu^A + K_{AB}P^B)\, dv^A + (K_{BA}\nu^B + I_{AB}P^B)\, d\Omega^A \quad .$$

Then the canonical form of the Euler-Lagrange equations (2.5 (19)) : $Z \lrcorner\, \omega - dE = 0$ implies

(14) $$K_{AB}\Phi^B + m\xi_A - \epsilon_{ABC}\, \Theta^B P^C + \frac{1}{2} \epsilon_{LMN}\, \Sigma^L R^{MN}{}_{AB}\, \nu^B = 0$$

(15) $$I_{AB}\dot{\Phi}^B + K_{BA}\dot{\xi}^B - \epsilon_{ABC}\,\Sigma^B P^C - \epsilon_{ABC}\,\Theta^B \nu^C = 0$$

(16) $$m\dot{\nu}_A + K_{AB}\dot{P}^B = m\dot{v}_A + K_{AB}\dot{\Omega}^B$$

(17) $$K_{BA}\nu^B + I_{AB}P^B = I_{AB}\Omega^B + K_{BA}\,v^B$$

using (10) and (13).

Equations (16) and (17) may be rewritten as

(18) $$m(\nu_A - v_A) + K_{AB}(P^B - \Omega^B) = 0 \quad \text{and}$$

(19) $$K_{BA}(\nu^B - v^B) + I_{AB}(P^B - \Omega^B) = 0 \quad ;$$

then the substitution of (18) into (19) gives

(20) $$(m\,I_{AB} - K_{CA}K^C{}_B)(P^B - \Omega^B) = 0\ .$$

If $||K|| < \sqrt{m||I||}$ in any suitable matrix norm $||\cdot||$, then $mI - K^tK$ has an inverse, which we denote by $\bar{I}$.

We conclude from (20) that

(21) $$P^A = \Omega^A \quad \text{and (18) gives}$$

(22) $$\nu^A = v^A\ .$$

To save some writing, set

(23) $$R_A = \epsilon_{ABC}\Theta^B\Omega^C - \frac{1}{2}\,\epsilon_{LMN}\,\Sigma^L R^{MN}{}_{AB}v^B \quad \text{and}$$

(24) $$\overline{R}_A = \epsilon_{ABC} \Sigma^B \Omega^C + \epsilon_{ABC} \Theta^B v^C \quad .$$

Substituting (14) into (15) and using (21) to (24) gives

(25) $$\Phi^A = \overline{I}^{AB}(m\overline{R}_B - R^C K_{CB}) \quad \text{and}$$

(26) $$\xi^A = m^{-1}(R^A - mK_{AB}\overline{I}^{BC}\overline{R}_C + K_{AB}\overline{I}^{BC} I_{CD} R^D)$$

where again $\overline{I}^{AB}$ denotes the inverse to

$$m\, I_{AB} - K_{CA} K^C{}_B \quad .$$

We have essentially solved the problem, the components of the vector field Z being given by (21), (22), (25) and (26). What is of interest, however, is the motion of the centre of mass in space, the time development of angular velocity, etc. This is achieved by projecting Z into $SO(\Sigma)$ and Σ . To do this let t denote the parameter of the integral curve of Z ; physically, t represents time.

The tangent vector of the trajectory of centre of mass is

(27) $$v^\Gamma \equiv \frac{dx^\Gamma}{dt} = Z \lrcorner dx^\Gamma = Z \lrcorner e^\Gamma_A \theta^A = \nu^A e^\Gamma_A \quad .$$

From (22) we see that

(28) $$v^\Gamma = v^A e^\Gamma_A \quad .$$

Similarly,

$$\frac{de^\Gamma_A}{dt} = Z \lrcorner de^\Gamma_A$$

$$= Z\lrcorner(-e_K\epsilon^K{}_{AL}\tilde{F}^L - \Gamma^\Gamma{}_{\wedge\Delta}e^\Delta_K e^\wedge_A\theta^K) \quad , \quad \text{so}$$

(29) $$\dot{e}^\Gamma_A = -e^\Gamma_K\epsilon^K{}_{AL}\Omega^L$$

where "$\cdot$" $= v^\Gamma\nabla_\Gamma$ is covariant differentiation along the trajectory $t \to x^\Gamma(t)$.

Next,

(30) $$\frac{d\Omega^A}{dt} = Z\lrcorner\, d\Omega^A = \Phi^A$$

so that computing Φ^A with the aid of (23) and (24) we get

(31) $$\frac{d\Omega^A}{dt} = \bar{I}^{AB}\{m\,\epsilon_{BKL}(\Sigma^K\Omega^L+\Theta^K v^L)-(\epsilon^C{}_{KL}\Theta^K\Omega^L - \frac{1}{2}\,\epsilon_{LMN}\Sigma^L R^{MNC}{}_K v^K)K_{CB}\} \quad .$$

For the linear acceleration, we have

$$\frac{dv^\Gamma}{dt} = Z\lrcorner\, d(e^\Gamma_A v^A) = v^A\, Z\lrcorner\, de^\Gamma_A + e^\Gamma_A\, Z\lrcorner\, dv^A \quad ,$$

or using (28), (29) and (26)

(32) $$\dot{v}^\Gamma = -e^\Gamma_A\,\epsilon^A{}_{BC}v^B\Omega^C$$
$$+ m^{-1}\, e^\Gamma_A\{\epsilon^A{}_{BC}\Theta^B\Omega^C - \frac{1}{2}\,\epsilon_{LMN}\Sigma^L R^{MNAB}v_B - mK^A{}_B\bar{I}^{BC}[\epsilon_{CKL}(\Sigma^K\Omega^L+\Theta^K v^L)$$
$$+ I_{CD}(\epsilon^D{}_{KL}\Theta^K\Omega^L - \frac{1}{2}\,\epsilon_{LMN}\Sigma^L R^{MNDK}v_K)]\} \quad .$$

(e) Subcases of Case (i) include: free motion of the body where K_{AB} is taken to be zero in (d); (31) and (32) become respectively

(33) $$\frac{d\Omega^A}{dt} = -\tilde{I}^{AB}\, S_{BC}\, \Omega^C$$

(34) $$\dot{v}^{\Gamma} = m^{-1}\, S^{\Lambda\Delta}\, R_{\Lambda\Delta\Sigma}{}^{\Gamma}\, v^{\Sigma} \quad .$$ 4)

The other subcase is where

(35) $$K_{AB} \equiv m\epsilon_{ABC}\, \eta^C ,$$

η^C being a constant vector (in the body frame) which labels a point of the body other than the centre of mass. It then turns out that $m^{-1}\hat{I} = I - m^{-1}K^tK$ is the moment of inertia tensor with respect to the point $e^{\Gamma}_A\, \xi^A + x^{\Gamma}$ of the rigid body (compare formula 2.2(15)). Thus (d) with (35) is the description of the motion of the rigid body with respect to a point of the body labelled by η relative to the centre of mass.

(f) Case (ii): $L = \frac{1}{2}\, m\, \delta_{AB} v^A v^B + \frac{1}{2}\, I_{AB}\Omega^A\Omega^B + m\, A_A v^A + m\, B_A \Omega^A - m\phi$ is the Lagrangian (5). In succession we compute that

(36) $$\Theta_A = mv_A + mA_A$$

(37) $$\Sigma_A = S_A + mB_A$$

(38) $$E = \frac{1}{2}\, m\, v_A v^A + \frac{1}{2}\, I_{AB}\Omega^A\Omega^B + m\phi$$

(39) $$\left\{ \begin{array}{l} dE = m\, v_A dv^A + S_A\, d\Omega^A + m\phi_A\theta^A \\[2ex] \phi_A \equiv e^{\Gamma}_A \dfrac{\partial\phi}{\partial x^{\Gamma}} \end{array} \right.$$

$$(40)\quad \begin{cases} dA_A = \epsilon_{ABC} A^B \tilde{E}^C + A_{A|B}\ \theta^B \\ \\ A_{A|B} \equiv e^{\Gamma}_{A}\ A_{\Gamma|\wedge}\ e^{\wedge}_{B} \end{cases}$$

$$(41)\quad \theta_L = (mv_A + mA_A)\theta^A + (S_A + mB_A)\tilde{E}^A$$

so that

$$\omega = -d\theta_L \quad \text{, or, using (38) to (41)}$$

$$(42)\quad \omega = mv^A\ \epsilon_{ABC}\tilde{E}^B \wedge \theta^C - mdv_A \wedge \theta^A + mA_{[A|B]}\ \theta^A \wedge \theta^B$$

$$- \frac{1}{2} m\ B^A\ \epsilon_{ABC}\tilde{E}^B \wedge \tilde{E}^C + \frac{1}{2} S_{AB}(\tilde{E}^B \wedge \tilde{E}^C + \frac{1}{2} R^{AB}{}_{KL}\theta^K \wedge \theta^L)$$

$$- I_{AB}\ d\Omega^A \wedge \tilde{E}^B + mB_{A|B}\ \theta^A \wedge \tilde{E}^B \quad .$$

If again $Z = \nu^A \underset{\sim}{e}_A + P^A \underset{\sim}{E}_A + \xi^A \dfrac{\partial}{\partial v^A} + \Phi^A \dfrac{\partial}{\partial \Omega^A}$ denotes a typical vector field in $\mathfrak{X}(TSO(\Sigma))$,

$$(43)\quad Z \lrcorner \omega = -(m\xi_A + m\ \epsilon_{ABC}P^B v^C + 2mA_{[A|B]}{}^B + MB_{A|B}P^B - \frac{1}{2} S^{MN}R_{MNKA}\nu^K)\theta^A$$

$$- (I_{AB}\Phi^B + S_{AB}P^B - m\ \epsilon_{ABC}v^B \nu^C - m\nu^B B_{B|A} - m\ \epsilon_{ABC}\ \Omega^B B^C)\tilde{E}^A$$

$$+ m\nu_A\ dv^A + I_{AB}P^B\ d\Omega^A \quad .$$

The equation $Z \lrcorner \omega = dE$ holds iff

$$(44)\quad \nu^A = v^A$$

$$(45)\quad P^A = \Omega^A$$

(46) $$\xi^A = \epsilon^A{}_{BC} v^B \Omega^C - 2\, A^{[A|B]} v_B - B{}_{|B}\, \Omega^B + \frac{1}{2} S^{MN} R_{MNK}\, v^K - \phi$$

(47) $$I_{AB} \dot{\Omega}^B = -S_{AB} \Omega^B + m v^B\, B_{B|A} + m\, \epsilon_{ABC}\, \Omega^B B^C \quad .$$

As an example, we can write the equations determining an integral curve of Z in three-dimensional vector notation for the special case that $\gamma = \delta_{\Gamma\Lambda}\, dx^\Gamma \otimes dx^\Lambda$ and $\vec{B}$ is constant with $\vec{B} = \nabla \times \vec{A}$:

(48) $$\dot{\vec{v}} = \vec{v} \times \vec{B} - \nabla\phi$$

(49) $$\dot{\vec{S}} = \vec{\Omega} \times \vec{S} + \vec{\Omega} \times \vec{B}$$

where $\vec{v} = (v^\Gamma)$ and $\vec{S} = (S_A)$. It is evident that (48) is an analogue of the usual Lorentz-force, and that (49) represents some sort of interaction of spin with an external magnetic field $\vec{B}$.

CHAPTER IV

Motion in Curved Space-Time

§4.1 Introduction.

So far, our discussions of the rigid body have taken the form of describing the evolution of the relevant physical quantities such as position of centre of mass and rotational configuration, dependent upon a (fixed) parameter, the time. As the title of this chapter indicates, we now generalize the formalism to include time as one of the coordinates; the motion will then appear as a (unparameterized) curve in a suitable (extended) configuration space. In doing so, we can not only discuss the relativistic rigid test body, but also include naturally a parallel description based on Galilean relativity.

Although we still rely on a definition of rigidity, such as 2.2 (1), in a nonrelativistic setting, the corresponding definition in general relativity is not immediately obvious (an extended body cannot be considered at one time, since simultaneity can't be defined over a finite region of space). Any "reasonable" rigidity condition, such as Born rigidity, for an extended body seems to be too restricting ([41]). However, we may retain some of the properties of a Newtonian rigid body if we adopt Dixon's dynamical rigidity condition ([13]) which essentially says that the classical relation (for example, 2.4 (23)) between spin angular momentum and angular velocity is valid for the body. In this case, there exists a mass constant, m, such that the rest energy, M, of the body has the expression

$$\text{(1)} \qquad M = m + T_{rot} + \left(\begin{array}{c}\text{potential energy of interaction}\\ \text{with applied fields}\end{array}\right)$$

We place an additional assumption on the dynamically rigid body, namely that quadrupole and higher multipole interactions with the gravitational field are negligible. Then the last term in (1) does not appear when there are no electromagnetic fields present.

In the relativistic case then, we confine our discussion to a rigid test body which is assumed to be characterized by the quantities defined in §2.4. Our model (in both the relativistic and Newtonian cases) consists of an orthonormal triad of space-like vectors fixed in the body, determining the orientation of the whole body at a given point in space-time. Hence we think of a rigid body existing, at a given event, only in the tangent space to space-time at this event. In practice, the model represents, for example, a gyroscope orbiting the earth.

With these understandings, we begin in the first three sections by reviewing the Galilei and Lorentz groups and algebras, Galilei and Lorentz structures on space-time, and the general programme of symplectic mechanics, which will be the basis of the model. Here, we are mainly interested in establishing the notation. Discussions of the Lorentz group can be found almost everywhere while the Galilei group and Galilean relativity is extensively reviewed in Levy-Leblond [30] and treated in several recent books [42], [52]. Galilei structures and corresponding comparisons of Galilean and Lorentz physics can be found in Havas [18] (traditional), Trautman [55] and Künzle [28], for example.

The last two sections motivate and describe the evolution space, presymplectic form, its kernel, and the space-time description of the motion. The treatment is, for the most part, that in Künzle [28], although the Galilean treatment is new. We refrain from discussing symmetries and conserved quantities since this is a deep topic in itself, and would lead us too far astray to give an adequate treatment here. (However, see [26] and [28], for example.)

§4.2 Galilei and Lorentz Groups and Their Algebras.

(a) Galilei group. Let (t,x) denote a point of R^4 with $x = (x^1,x^2,x^3) = (x^A)$. The proper Galilei group, $\widetilde{\mathrm{Gal}}$, can be defined as the group of affine endomorphisms of R^4 of the form

$$(1)\quad \left\{ \begin{array}{l} (t,x) \to (t',x') = (t+b, Rx+vt+a) \quad , \\ \\ b \in R \ , \ a,v \in R^3 \ , \ R \in SO(3) \quad . \end{array} \right.$$

If $v = 0$, we have essentially the Euclidean group of R^3 acting on the (affine) subspace $t = \text{const.}$ For $b = 0$, $a = 0$, and $R = \text{id}$, we obtain the pure Galilean tranformations which are interpreted as switching on a motion with velocity v , <u>effected at time</u> $t = 0$ in the above parameterization. Corresponding to $a = 0$ and $b = 0$ is the homogeneous Galilei group, Gal .

Instead of considering Gal as an endomorphism group of R^4 , we will mainly think of Gal as operating on Galilean frames (to be made precise shortly) where (t,x) , (t',x') are the coordinates of

the same point of R^4 in two Galilean frames.

(b) Lorentz group. Here we consider the connected component of the identity only, and denote it by Lor. We can represent an element $L \in \mathrm{Lor}$ by the matrix

$$(2)\quad \begin{cases} L = \left[\exp\begin{pmatrix} 0 & b \\ b^t & 0 \end{pmatrix}\right]\begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} \\ \\ A \in SO(3) \quad , \quad b \in R^3 \quad , \end{cases}$$

and where exp denotes the exponential map to a Lie group from its algebra (cf. §2.3).

(c) For comparison purposes, we represent Gal and Lor in $GL(4)$ acting on the left on R^4 :

$$(3)\quad \begin{cases} a \in \mathrm{Gal} \leftrightarrow a = \begin{pmatrix} 1 & 0 \\ b & R \end{pmatrix} \\ \\ L \in \mathrm{Lor} \leftrightarrow L = \begin{pmatrix} \alpha & c^{-2}R^t b \\ b & R + \left(\frac{\alpha - 1}{b^2}\right) b(R^t b) \end{pmatrix} \\ \\ \qquad b \in R^3 \; , \; R \in SO(3) \\ \\ \alpha = (1 + c^{-2}b^2)^{1/2} \qquad c : \text{speed of light.} \end{cases}$$

It is clear that $\lim\limits_{c\to\infty} L = a$.

(d) Lie algebras. Throughout the sequel we define [5)]

(4) $$\nu = \begin{cases} 0 \quad , & \text{Galilei case} \\ 1 \quad , & \text{Lorentz case} \quad . \end{cases}$$

Let $\{E_a^b\}$ denote the standard basis of $gl(4)$. Take, then, as bases of the Lorentz algebra, lor, and Galilei algebra, gal, the matrices

(5) $$\begin{cases} \{E_{AB}, F_A\} \\ E_{AB} = E_B^A - E_A^B \quad , \quad 1 \leq A < B \leq 3 \\ F_A = -E_A^o - \nu E_o^A \quad . \end{cases}$$

Noting that $\{E_{AB}\}$ generates $so(3)$, hence $\{E_A = -\frac{1}{2}\epsilon_A{}^{BC} E_{BC}\}$ also generates $so(3)$, we often will use $\{E_A, F_A\}$ as a basis of gal $(\nu=0)$ or lor $(\nu=1)$.

§4.3 Galilei and Lorentz Structures on Space-Time.

(a) Action of Gal on frames of a 4-dimensional vector space V . Let

(1) $$\begin{cases} (e_a) & \text{be a frame (basis) of } V \text{ , and} \\ (\theta^a) & \text{the dual frame of } V^* = L(V,R) \text{ , so that} \\ \theta^a(e_b) = \delta_b^a \quad . \end{cases}$$

Let $x \in V$, $x = x^a e_a$. Then, if $G \in Gal$, the coordinate transformation $x^a \to \tilde{x}^a = G^a{}_b x^b$ corresponds to a basis transformation $e_a = \tilde{e}_b G^b{}_a$ of

V , and the transformation of the dual basis of V^* , $\tilde{\theta}^a = G^a{}_b\, \theta^b$.

Recalling 4.2 (3) we have

$$(2) \qquad \begin{cases} G^o{}_o = 1 & G^c{}_A = 0 \\ G^A{}_o = b^A & G^A{}_B = R^A{}_B \\ (b^A) \in R^3 \ , & (R^A{}_B) \in SO(3) \ , \end{cases}$$

so that the basis transformations are written as:

$$(3) \qquad \begin{cases} \tilde{e}_o = e_o - e_A (R^{-1})^A{}_B\, b^B \\ \tilde{e}_A = e_B (R^{-1})^B{}_A \\ \tilde{\theta}^o = \theta^o \\ \tilde{\theta}^A = b^A \theta^o + R^A{}_B\, \theta^B \quad . \end{cases}$$

These relations (3) imply that the tensors

$$\psi = \theta^o \quad \text{and} \quad \gamma = \delta^{AB}\, e_A \otimes e_B$$

are invariant under the action of Gal ; conversely, given a 1-form, $\psi \in V^*$, and a symmetric rank-3 contravariant tensor, $\gamma \in V \otimes V$, such that $\gamma(\psi) = 0$, we can characterize an action of Gal on V .

In the following remarks, $\Lambda = (\Lambda^a{}_b)$ denotes a generic element of either Gal or Lor.

(b) Bundle of {Gal | Lor} frames. We have defined the bundle of linear frames in the Appendix. The strategem now is to restrict ourselves to those frames (at a point) related by a {Gal | Lor} transfor-

mation (compare §3.26); this we make precise.

Let V be a (C^∞-) four dimensional manifold (space-time). Suppose that at each point $x \varepsilon V$ we are given a frame $\underset{o}{e} = \underset{o}{e}_r \varepsilon L_x V$, and that this assignment, $x \to \underset{o}{e}(x)$, is smooth. Let G denote either of the groups Gal or Lor.

For each $x \varepsilon V$, define $G_x V$ by:

$$(4) \qquad e = (e_a) \varepsilon G_x V \quad \text{iff} \quad e_a = e_{or} \Lambda^r{}_a \quad \text{for some} \quad \Lambda \varepsilon G .$$

Then set $G(V) = \bigcup_{x \varepsilon V} G_x V$. $G(V)$ is a principal fibre bundle over space-time V with group $G = \{\mathrm{Gal} \mid \mathrm{Lor}\}$. Note that $G(V)$ depends on the choice of $x \to \underset{o}{e}(x)$. V together with such a <u>reduction</u> of $L(V)$ to $G(V)$ is called a G <u>manifold</u>.

(c) Tensor characterizations of $G(V)$. Now we must separate the two cases according to whether $G = \mathrm{Gal}$ or $G = \mathrm{Lor}$. Consider $\mathrm{Lor}(V)$ as in (b) and denote by η_{ab} the matrix, diag$(-1,1,1,1)$. Define the smooth symmetric tensor field on V $g = g_{\alpha\beta}\, dx^\alpha \otimes dx^\beta$ by

$$(5) \qquad g_{\alpha\beta}(x) = \theta^a_{o\alpha}(x)\, \eta_{ab}\, \theta^b_{o\beta}(x)$$

where, as usual, $\underset{o}{\theta}^a = \theta^a_{o\alpha}\, dx^\alpha$ denotes the frame of $T^*_x V$ dual to $\underset{o}{e}$. Clearly, this tensor field, $g = g_{\alpha\beta}\, dx^\alpha \otimes dx^\beta = \eta_{ab}\, \theta^a \otimes \theta^b$, is invariant by Lor. Conversely, given a Lorentz metric, $g = g_{\alpha\beta}\, dx^\alpha \otimes dx^\beta$, on V, a Lorentz structure is determined by: $(x, e_a) \varepsilon \mathrm{Lor}(V)$ iff $e^\alpha_a\, g_{\alpha\beta}(x)\, e^\beta_b = \eta_{ab}$.

Recalling the results of (a), we can similarly define a tensor characterization of $Gal(V)$. Assume that we are given a Galilean structure on V as in (b). Define the following 1-form, $\psi = \psi_\alpha \, dx^\alpha$, and symmetric rank-3 tensor field, $\gamma = \gamma^{\alpha\beta} \, \partial_\alpha \otimes \partial_\beta$:

$$(6) \qquad \begin{cases} \psi_\alpha(x) = \theta^{o}_{o\alpha}(x) \\ \gamma^{\alpha\beta}(x) = e^{\alpha}_{oA}(x) \; \delta^{AB} \; e^{\beta}_{oB}(x) \quad . \end{cases}$$

Conversely, suppose that we are given a 1-form $\psi = \psi_\alpha \, dx^\alpha$ and a symmetric rank-3 tensor field $\gamma = \gamma^{\alpha\beta} \, \partial_\alpha \otimes \partial_\beta$ such that $\gamma^{\alpha\beta} \psi_\beta = 0$; define a Galilean structure on V by:

$$(x,e_a) \in Gal(V) \iff \begin{cases} e^{\alpha}_{a} \, \psi_\alpha = \delta^{o}_{a} \qquad \text{and} \\ \theta^{a}_{\alpha} \, \gamma^{\alpha\beta} \, \theta^{b}_{\beta} = \delta^{a}_{A} \, \delta^{AB} \, \delta^{b}_{B} \quad . \end{cases}$$

(d) Terminology. Let V , $G(V)$ be as above, and suppose $x \in V$. A vector $\xi = \xi^\alpha \, \partial_\alpha \in T_x V$ is called

<u>timelike</u> iff $\begin{cases} \psi_\alpha \, \xi^\alpha \neq 0 \, , & \text{Gal} \\ g_{\alpha\beta} \, \xi^\alpha \xi^\beta < 0 \, , & \text{Lor} \end{cases}$

<u>spacelike</u> iff $\begin{cases} \psi_\alpha \, \xi^\alpha = 0 \, , & \text{Gal} \\ g_{\alpha\beta} \, \xi^\alpha \xi^\beta > 0 \, , & \text{Lor} \end{cases}$

and, in the Lorentz case only,

<u>null</u> iff $\quad g_{\alpha\beta} \, \xi^\alpha \xi^\beta = 0$.

Note that in the Galilei case ψ defines a 3-dimensional distribution, $x \to S_x \subset T_x V$, defined by: $\xi \in S_x$ iff $\xi \lrcorner \psi = 0$; that is, S_x consists of all the spacelike vectors of $T_x V$. If $x \to S_x$ is integrable (i.e., the Lie bracket, $[\xi,\eta]$, of any two vectors, ξ, η, in S_x is again in S_x), then each of the submanifolds $\{x \,;\, x^o = t\}$ carries an induced Riemannian metric $(,)$: if $\xi = \xi^A e_A$ and $\eta = \eta^A e_A \in S_x$, e_a being any Gal frame at x, then $(\xi,\eta) = \xi^A \delta_{AB} \eta^B$.

We will be concerned with only those Galilean manifolds, (V,ψ,γ), which are <u>first order flat</u>[6]; that is, those for which ψ satisfies

$$\text{(7)} \qquad d\psi = 0 \quad .$$

(e) Lor-connections. We do not need to say much about the connection which we use on $Lor(V)$: it is merely the Levi-Civita connection induced by the metric g. The lor-valued connection 1-form is

$$\text{(8)} \quad \left\{ \begin{array}{l} \omega = \tilde{E}^A E_A + \tilde{F}^A F_A \;, \quad \text{where} \\ \tilde{E}^A = -\frac{1}{2} \epsilon^{A}{}_{B}{}^{C} \theta^B_\alpha (de^\alpha_C + \Gamma^\alpha_{\beta\gamma} e^\gamma_C dx^\beta) \\ \tilde{F}^A = \theta^A_\alpha (de^\alpha_o + \Gamma^\alpha_{\beta\gamma} e^\gamma_o dx^\beta) \\ \text{and the } \Gamma^\alpha_{\beta\gamma} \text{ are the usual Christoffel symbols.} \end{array} \right.$$

(f) Gal-connections. The situation now is a bit more arbitrary than that in the Lorentz case. (γ,ψ) does not determine a symmetric torsion-free Galilei connection uniquely, but it turns out that ([28])

if (V,ψ,γ) is a first order flat (f.o.f.) Galilei manifold, any connection is torsion-free and the set of all connections on V is in a 1-1 correspondence with the set of all 2-forms on V.

In a f.o.f. Galilei manifold, (V,ψ,γ), we have $d\psi = 0$, so that (locally) $\psi = t : V \to R$; hence we can find (locally) an adapted coordinate system such that

$$(9)\qquad \begin{cases} \psi_\alpha = \delta^o_\alpha \\ \gamma^{\alpha o} = 0 \\ (\gamma_{\Gamma\Delta}) = (\gamma^{\Gamma\Delta})^{-1} \end{cases}$$

The components, in this coordinate system, of the most general Galilei connection are

$$(10)\qquad \begin{cases} \Gamma^o_{\alpha\beta} = 0 \\ \Gamma^\Gamma_{oo} = 2\delta^{\Gamma\Lambda}\,\kappa_{o\Lambda} \\ \Gamma^\Gamma_{o\Lambda} = \delta^{\Gamma\Delta}(\kappa_{\Lambda\Delta} - \frac{1}{2}\,\partial_o\,\gamma_{\Lambda\Delta}) \\ \Gamma^\Gamma_{\Lambda\Delta} = \frac{1}{2}\,\gamma^{\Gamma\Sigma}(-\partial_\Sigma\gamma_{\Lambda\Delta} + 2\partial_{(\Lambda}\gamma_{\Delta)\Sigma}) \end{cases}.$$

The curvature tensor of a symmetric connection on a f.o.f. Galilei manifold satisfies:

$$(11)\qquad \begin{cases} R^\alpha{}_{[\beta\gamma\delta]} = 0 \\ R^\alpha{}_{\beta[\gamma\delta,\rho]} = 0 \\ \psi_\rho\, R^\rho{}_{\beta\gamma\delta} = 0 \\ \gamma^{\rho\beta}\, R^\alpha{}_{\rho\gamma\delta} = 0 \end{cases}$$

If, in addition, we require that

$$\gamma^{\rho[\alpha}R^{\gamma]}{}_{(\beta\delta)\rho} = 0 \quad , \tag{12}$$

then there are enough algebraic restrictions on $R^{\alpha}{}_{\beta\gamma\delta}$ to reduce the number of independent components to 20 as for the Lorentz case. Equation (12), in fact, is equivalent to saying that on a f.o.f. Galilei manifold the connection satisfies

$$d\kappa = 0 \tag{13}$$

in the notation of (9) and (10). (This is expressed more generally in [28] than we need here.) Such a connection has been called a Newtonian connection ([28]).

(g) General formulae for reference. Here V is any Lorentz manifold (V,g) or f.o.f. Galilei manifold (V,ψ,γ) with Newtonian connection. As usual,[5)] $\nu = 0$ refers to the Galilei case, $\nu = 1$ to the Lorentz, while G refers to one of Gal or Lor , g to gal or lor .

The g - valued connection form with respect to the basis $\{E^a{}_b\}$ of $gl(4)$ is $\omega = \omega^a{}_b\, E_a{}^b$; also,

$$\left\{\begin{array}{l} \omega = \tilde{E}^A E_A + \tilde{F}^A F_A \quad \text{where} \\ \omega^A{}_B = -\epsilon^A{}_{BC}\, \tilde{E}^C \\ \omega^o{}_A = \nu\, \tilde{F}_A \\ \omega^A{}_o = \tilde{F}^A \quad . \end{array}\right. \tag{14}$$

(Exterior) differentials of tensors expressed in G - frame components:

Denote

(15) $$T^{a\ldots b}_{k\ldots \ell} \equiv \theta^a_\alpha \ldots \theta^b_\beta \, e^\rho_k \ldots e^\sigma_\ell \, T^{\alpha\ldots\beta}_{\rho\ldots\sigma} \quad .$$

Next,

(16) $$\begin{cases} de^\alpha_o = e^\alpha_K \tilde{F}^K - \Gamma^\alpha_{\beta\gamma} \, e^\beta_r \, e^\gamma_o \, \theta^r \\ de^\alpha_A = -e^\alpha_K \, \epsilon^K{}_{AL} \, \tilde{E}^L + \nu \, e^\alpha_o \, \tilde{F}_A - \Gamma^\alpha_{\beta\gamma} \, e^\beta_r \, e^\gamma_A \, \theta^r \end{cases}$$

while

(17) $$\begin{aligned} & d\, T^{a\ldots b}_{k \quad \ell} + \omega^a{}_r \, T^{r\ldots b}_{k\ldots\ell} + \ldots + \omega^b{}_r \, T^{a\ldots r}_{k\ldots\ell} \\ & -\omega^r{}_k \, T^{a\ldots b}_{r\ldots\ell} - \ldots - \omega^r{}_\ell \, T^{a\ldots b}_{k\ldots r} = T^{a\ldots b}_{k\ldots\ell|r} \, \theta^r \quad . \end{aligned}$$

The structure equations of a G - connection with no torsion, and curvature 2-form $\Omega^a_b = \frac{1}{2} R^a{}_{brs} \; \theta^r \wedge \theta^s$ are:

(18) $$\begin{cases} d\theta^o = \nu \, \theta_A \wedge \tilde{F}^A \\ d\theta^A = -\epsilon^A{}_{BC} \, \theta^B \wedge \tilde{E}^C - \tilde{F}^A \wedge \theta^o \\ d\tilde{E}^A = \frac{1}{2} \epsilon^A{}_{BC} (-\tilde{E}^B \wedge \tilde{E}^C + \nu \, \tilde{F}^B \wedge \tilde{F}^C - \frac{1}{2} R^{BC}{}_{rs} \, \theta^r \wedge \theta^s) \\ d\tilde{F}^A = -\epsilon^A{}_{BC} \, \tilde{E}^B \wedge \tilde{F}^C + \frac{1}{2} R^A{}_{ors} \, \theta^r \wedge \theta^s \quad . \end{cases}$$

§4.4 The General Programme of Symplectic Mechanics.

(a) In view of the increasing popularity of symplectic mechanics and the number of diverse forms the theory is being developed in, it may well be useful to include a brief survey of the field and to make explicit the form that we use in the sequel. Without going into detail, we hope to resolve some of the ambiguity of the varying terminology and usage of the various structures found in the literature.

(b) Inhomogeneous canonical dynamics. This formalism (developed in [15], [51], and especially Abraham [1]) was discussed in some detail in §2.5. Suppose Q (a C^∞ n - dimensional manifold) is the configuration space of some system; we saw that T^*Q has a canonical symplectic structure, $\omega_o = -d\theta_o$, where θ_o is the canonical form on T^*Q. $M = T^*Q$ is known as the momentum phase space or simply phase space. A vector field on M $X \in \mathfrak{X}(M)$ is called locally Hamiltonian iff $L_X \omega_o = 0$ (X generates a symmetry of ω_o.) Then locally there is a function H such that $X = X_H$ where X_H is defined by $X_H \lrcorner\, \omega_o = dH$. H is called a (local) Hamiltonian for X_H. Hamilton's equations of motion are essentially the equations for the integral curves of X_H.

In §2.5 we were given a suitable function $L : TQ \to R$, the Lagrangian of a mechanical system. TQ is the velocity phase space or just state space. In terms of L, we defined a (regular) map $FL : TQ \to T^*Q$ (the fibre derivative) and an energy function $E : TQ \to R$. Via FL, the symplectic structure is transferred to TQ, and the canonical form of Lagrange's equations obtained as integral curves of the vector field defined by $X_E \lrcorner\, \omega = dE$. The projections of these curves

to Q coincide with those obtained from Hamilton's equations when $H : T^*Q \to R$ is derived from E via FL (that is, $H = E \circ (FL)^{-1}$) if FL is a diffeomorphism.

(c) Contact Manifolds. An odd-dimensional manifold M on which is defined a closed 2-form ω of maximal rank is called a <u>contact manifold</u> and ω a <u>contact structure</u>.[7)] Contact structures are basic to time-dependent mechanics and can describe either Hamiltonian systems or Lagrangian systems. We give a simple example of the former - Lagrangian systems are similar.

Let (M,ω) be a symplectic manifold and let $pr : R \times M \to M$ be the projection. For a given regular function $H : R \times M \to R$, define the 2-form

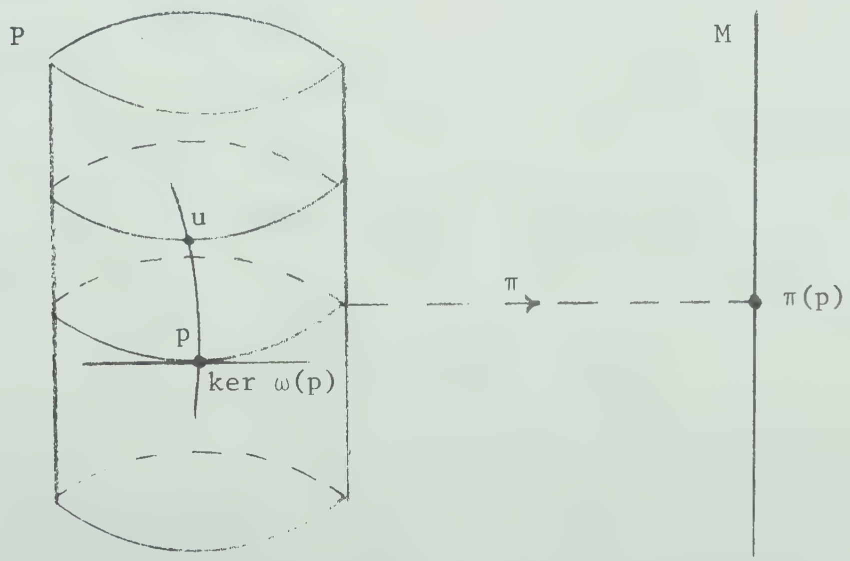
$$\tilde{\omega} = pr^*\omega + dH \wedge dt \quad . \tag{1}$$

Then $(R \times M, \tilde{\omega})$ is a contact manifold ([1]). A motion of the system is a nonparametrized integral curve of a vector field satisfying $X \lrcorner\, \tilde{\omega} = 0$. If it happens that $\omega = d\theta$ then

$$\begin{aligned} \tilde{\omega} &= d\tilde{\theta} \\ \tilde{\theta} &= pr^*\theta + H\,dt \quad . \end{aligned} \tag{2}$$

(d) Evolution space; space of motions. Denoted by P and M respectively, these were introduced by Souriau (see for example, [48]). Evolution space P is a generalization of the contact manifolds of (b); for a Hamiltonian expressed in homogeneous coordinates (time is included as one of the coordinates), $P = T^*(R \times Q)$. $P = R \times T^*Q$ is the evolution

space of a (Hamiltonian) contact manifold.[8)] Defined in such an evolution space is a closed 2-form ω of constant rank $< \dim P$, that is, a pre-symplectic form. ω defines a characteristic distribution $p \to \ker \omega(p) \subset T_pP$ which is integrable since $d\omega = 0$, and thus $p \to \ker \omega(p)$ is a foliation. Associated to this characteristic foliation is the quotient set $M = P/\ker \omega$ of leaves (that is, maximal integral submanifolds of $p \to \ker \omega(p)$). Denote the projection $P \to M$ by π - to each point $p \in P$, π associates the (unique) leaf passing through p. By requiring that π be both open and continuous, a unique topology on M is defined. On the other hand, the existence of a differentiable structure on M such that M becomes a quotient manifold and $\pi : P \to M$ a submersion, is not as easy. Call a submanifold U of P containing a point $p \in P$ transversal to the foliation $\ker \omega$ if at each point $q \in U$, T_qU is supplementary to $\ker \omega(p)$; that is, $T_qP = T_qU \oplus \ker \omega(p)$. U is a section of $\ker \omega$ iff U meets each leaf of the foliation in at most one point (or not at all).



Then, the foliation $\ker \omega$ is said to be <u>sectile</u> (or <u>regular</u>) iff for each $p \in P$ there is a submanifold $U \subset P$ such that $p \in U$, U is a section of $\ker \omega$, and U is transversal to $\ker \omega$.

<u>Theorem</u>: With the above definitions, $M = P/\ker \omega$ is a quotient manifold iff $\ker \omega$ is regular. (For a proof, see Palais [37], page 19.)

It the theorem holds, there exists a unique symplectic form $\tilde{\omega}$ on M such that $\omega = \pi^{*}\tilde{\omega}$; that is, ω on P is the pull-back of $\tilde{\omega}$ on M. M is the <u>space of motions</u>; each point of M is a leaf of $\ker \omega$, that is, a maximal connected integral manifold of the distribution $p \to \ker \omega(p)$. In effect (P,ω) is analogous to the example described in (b): ω describes the dynamics on P ; one solves for (local) vector fields Z on P satisfying $Z \lrcorner\, \omega = 0$ (iff $Z \in \ker \omega$), and the solution manifolds are the leaves of $\ker \omega$.

$(M,\tilde{\omega})$ then contains all the canonical structure of the system, for example, the Poisson bracket structure on the space of all observables of M. An <u>observable</u> f on M is defined to be a real-valued (smooth) function on M, so we show how to define a Lie-algebra operation on $C^{\infty}(M)$. Suppose $f,g \in C^{\infty}(M)$; recall that $\tilde{\omega}$ defines an isomorphism $\phi_{\tilde{\omega}} : \mathfrak{X}(M) \to \mathfrak{X}^{*}(M)$ of vector fields to 1-forms: $\phi_{\omega}(Z) = Z \lrcorner\, \tilde{\omega} \in \mathfrak{X}^{*}(M)$ for $Z \in \mathfrak{X}(M)$. Then

$$\{f,g\} = (\phi_{\tilde{\omega}}^{-1}(df)) \lrcorner\, dg \tag{1}$$

is the <u>Poisson bracket</u> of f and g. Local expressions and further properties of Poisson brackets are found, for example, in [1]. Canonical structures of classical systems is of contemporary interest because of its

role in canonical quantization ([21], [10]). Although the theory of finite dimensional systems is very well developed as well as the main interest being field-theoretic, a thorough analysis of finite-dimensional systems is often indispensible in the study of infinite-dimensional situations. The latter are symplectic manifolds modelled on (i.e., coordinatized by) Banach spaces. A lot of the finite dimensional theory carries over unchanged, but care must be taken, for example, with the different topologies involved. For a recent survey, see Chernoff and Marsden [7].

§4.5 Space-Time Model of a Free Rigid Body.

(a) We now take the principal bundle of G - frames, P, over space-time V as evolution space for the rigid body (notation - §4.3). The canonical system (P,ω) is defined in this section, and the calculations of the canonical distribution $\ker \omega$ (§4.4c) and the associated equations of motion in space-time are carried out in the following section (§4.6). The treatment of the relativistic system given here is essentially that of [28] while the Galilean, in this formalism, appears to be new.

(b) Space-time. Let V denote a four-dimensional differentiable (C^{∞}-) manifold. The following properties of a space-time are not unreasonable:

(i) Topological: paracompact, Hausdorff, connected;

(ii) Orientability: V is oriented (i.e., there is a globally defined volume element σ on V);

(iii) (Time Orientability): there exists a tangent line bundle on V .

The assumptions (i) are not special; they appear to be standard in the literature. It may be pointed out that paracompactness implies the existence of, first, a partition of unity on V , and second, a (global) Riemannian metric on V ([59]).

Property (iii) is equivalent to the global existence of both a Galilei and a Lorentz structure on V (cf. Steenrod [50], page 207). Even if (iii) is not assumed, it will hold for any compact manifold whose Euler characteristic is zero ([50], page 203) or for any non-compact manifold ([61], Theorem 4.8). For more details, see [60].

(ii) and (iii) together in the relativistic case imply that V is space and time orientable. Physically, it appears that time-orientability implies space-orientability. This is argued on the basis of elementary particle reaction experiments and also the CPT-theorem (see Hawking and Ellis [19] for a discussion and references).

It may be noted that (iii) additionally is equivalent to the existence of a globally nonvanishing vector field on V .

In what follows V will denote space-time with additional structure:

(iv) a first-order-flat Galilei structure with a Newtonian connection Γ , or

(iv)' a Lorentz structure with the induced Levi-Civita connection.

A free motion of a test particle will be taken as geodesic motion in both cases. Here comes an essential difference between Newtonian and Lorentzian physics: The connection in general relativity describes only the gravitational force; in Newtonian space-time, the connection may include other force laws in addition to that of gravitation. Referring to 4.3 (10), we may have a conservative force (gravitational; electrostatic) in $\Gamma^{\Lambda}_{oo} = \gamma^{\Lambda\Gamma} \partial_{\Gamma}\phi$ for some potential ϕ, or an electromagnetic-type force: $\Gamma^{\Lambda}_{o\Gamma} = 2\gamma^{\Lambda\Delta} \partial_{[\Delta}A_{\Gamma]}$, for a "vector potential" A.

(c) Space-time configuration of a rigid body. The space-time position of the centre of mass of a rigid body is a point $x \in V$. As in §3.2, the orientation of the body is described by a triad of space-like orthonormal vectors (e_A). Orientability of V allows us to choose a unit time-like vector e_o at x.[9] In the Lorentz case, this is unique; however, the condition in the Galilei case for e_o to be time-like unit, $e_o \lrcorner\, \psi = 1$, has a three-dimensional (affine) subspace of T_xV at each x as solution. Uniqueness may be regained by requiring e_o to be the unit tangent vector to the worldline of the centre of mass, however this is somewhat unsatisfactory. Alternatively, a unit time-like vector field u on V may be assumed to exist and required to have the coordinates δ^{α}_{o} in an adapted coordinate system (4.3(9)); this assumption is actually built into the formulae 4.3(10) ([28]). In any case, we assume a configuration of the rigid body is represented by a point $p = (x, e_a)$ in the bundle of G - frames (§4.3).

It is clear that the choice of $G(V)$ as configuration space for the rigid body is appropriate. Not only do the physical quantities

have natural interpretations in terms of the geometrical structure (explained more fully below), but it is easy to demand some degree of covariance of the theory.

(d) $TG(V)$ as evolution space. Recall (Appendix) that $TG(V)$ is parallelizable; $\{\underset{\sim}{e}_a, \underset{\sim}{E}_A, \underset{\sim}{F}_A\}$ forms a global basis of $\mathfrak{X}(G(V))$ while $\{\theta^a, \tilde{E}^A, \tilde{F}^A\}$ is the dual basis of $\mathfrak{X}^*(G(V))$. It was also pointed out that $\underset{\sim}{E}_A$ generates $so(3)$ at each point in V , so $\Omega^A E_A$ is again angular velocity at $x \in V$ in the body frame. Denote coordinates on $TG(V)$ by $(x^\alpha, e^\alpha_a; v^a, \Omega^A, U^A)$.

To see what to do next, examine the canonical one-form $\tilde{\theta}$, discussed with respect to time-dependent canonical dynamics in §4.4b, for the particular case of a free rigid body (no external forces are present):

(1) $$\tilde{\theta} = \frac{1}{2}(mv_A v^A + I_{AB}\Omega^A\Omega^B)dt - mv_A\theta^A - I_{AB}\Omega^A\tilde{E}^B$$

(see 3.3(8) and 4.4(2)). We follow Künzle [29] by taking as Cartan form

(2) $$\theta = M\theta^o - I_{AB}\Omega^A \tilde{E}^B$$

where

(3) $$M = m + \frac{1}{2} I_{AB}\Omega^A\Omega^B = m + T_{rot}$$

in the relativistic case only (cf. §4.1). Here, m denotes the mass of the body, measured in a frame in which the body is at rest, both translationally and rotationally.

As a motivation for (2), we first remark that, by assumption, the model of rotational motion developed in §2.4 is valid at each space-time event; whence, we retain the last term of (1). To justify the translational part of θ, we recall that if the two-form ω is exact (as it is here), $\omega = d\theta$, then associated to a dynamical symmetry of the system, generated by a vector field ξ (so that $L_\xi \omega = 0$), is a constant of the motion, f_ξ, given by $f_\xi = -\xi \lrcorner \theta$. Now, it is widely accepted that $p = Me_o$ is the four-momentum of an extended rigid body ([12], [13], [29]). p can be considered as a generator of space-time translations in the p - direction; under such a symmetry, the corresponding constant of the motion is $f_p = p \lrcorner \theta = -Me_o \lrcorner \theta = -M^2$, as would be expected.

To complete the definition of the canonical system $(T\,Lor(V), \omega)$ in the relativistic case, we define

$$\omega = d\theta \tag{4}$$

where θ is given in (3). The calculation is carried out using equations 4.3 (18) (with $\nu = 1$) and results in

$$\begin{aligned}\omega = {} & \frac{1}{4} S^{MN} R_{MNab}\, \theta^a \wedge \theta^b + S_A\, d\Omega^A \wedge \Omega^o - M\tilde{F}^A \wedge \theta_A \\ & + \frac{1}{2} S_{AB}(\tilde{E}^A \wedge \tilde{E}^B - \tilde{F}^A \wedge \tilde{F}^B) + I_{AB}\, \tilde{E}^A \wedge d\Omega^B ,\end{aligned} \tag{5}$$

with the notation

$$(6)\qquad \begin{cases} S_A = I_{AB}\,\Omega^B \\ S_{AB} = \epsilon_{AB}{}^C\, S_C \ . \end{cases}$$

In this case, $S_A\,\tilde{E}^A$ is the spin angular momentum in the body frame (compare 2.4 (24)).

We return to the Newtonian case and attempt to mimic the above procedure. If we try to use (2) to define a Newtonian Cartan form θ, we immediately run into insurmountable difficulties. For, not only does M reduce to m (a constant), so $dM = 0$, but also, for the Galilei structure, $d\theta^o = 0$. On passing to $\omega = d\theta$, we lose all translational information, and the procedure therefore breaks down.

What we will do then, is to postulate a presymplectic form in the Galilei case. To give a reasonable justification for our choice, we find it convenient to consider simultaneously a contact structure on $E = R \times T(R^3 \times SO(3))$ which we coordinatize (globally) by $(t; x^\Gamma, e^\Gamma_A; v^\Gamma, \Omega^A)$, and a flat Galilei structure on R^4 which is defined by

$$(7)\qquad \begin{cases} \psi_\alpha = \delta^o_\alpha \\ \gamma^{\alpha\beta} = \delta^\alpha_A \delta^{AB} \delta^\beta_B \ . \end{cases}$$

The Cartan form on E is

$$(8)\qquad \theta = \frac{1}{2}\,(m v_\Gamma v^\Gamma + I_{AB}\Omega^A\Omega^B)\,dt - m v_\Gamma dx^\Gamma - S_A\,\tilde{E}^A \ ,$$

so that

$$\omega = mv_\Gamma dv^\Gamma \wedge dt + S_A \, d\Omega^A \wedge dt - m dv_\Gamma \wedge dx^\Gamma$$
$$- I_{AB} \, d\Omega^A \wedge \tilde{E}^B + \frac{1}{2} S_{AB} \, \tilde{E}^A \wedge \tilde{E}^B \quad , \tag{9}$$

where we have used 2.4 (18). Now, we are used to thinking of v^Γ as the velocity of a free motion, that is, $v^\Gamma = \frac{dx^\Gamma}{dt}$; otherwise written,

$$v^\Gamma \, dt = dx^\Gamma \quad . \tag{10}$$

The exterior derivative of (10) yields $dv^\wedge \wedge dt = 0$, so that the first term of (9) disappears and we are left with

$$\tilde{\omega} = -m \, dv_\Gamma \wedge dx^\Gamma + S_A \, d\Omega^A \wedge dt$$
$$-I_{AB} \, d\Omega^A \wedge \tilde{E}^B + \frac{1}{2} S_{AB} \, \tilde{E}^A \wedge \tilde{E}^B \quad . \tag{11}$$

Now consider R^4 with the Galilei structure defined by (7) and flat connection, $\Gamma = 0$. From 4.4 (18) we calculate that $de_o^\Gamma = e_A^\Gamma \tilde{F}^A$. Thinking of e_o^Γ as a transformation to the velocity along the motion, we make the following correspondence between E and $Gal(R^4)$:

$$\left\{ \begin{aligned} dx^\Gamma &\leftrightarrow e_A^\Gamma \, \theta^A \\ dv^\Gamma &\leftrightarrow e_A^\Gamma \, \tilde{F}^A \\ dt &\leftrightarrow \theta^o \quad . \end{aligned} \right. \tag{12}$$

Using (12) in (11), obtain

$$\text{(13)} \qquad \tilde{\omega} = m\, \theta_A \wedge \tilde{F}^A + S_A\, d\Omega^A \wedge \theta^o + I_{AB}\, \tilde{E}^A \wedge d\Omega^A + \frac{1}{2} S_{AB}\, \tilde{E}^A \wedge \tilde{E}^B \quad .$$

The choice for ω in the general Galilean situation will be the simplest one reducing to $\tilde{\omega}$ in the case of force-free motion and flat Galilei structure, as well as being consistent with the relativistic form (5).

Hence, the model for a test rigid body in space-time is the canonical system $(G(V),\omega)$ where ω in both the Lorentz cnd Galilei cases being written as

$$\text{(14)} \qquad \begin{cases} \omega = \frac{1}{4} S^{MN} R_{MNab}\, \theta^a \wedge \theta^b + M\theta_A \wedge \tilde{F}^A + S_A\, d\Omega^A \wedge \theta^o \\ \qquad + I_{AB}\, \tilde{E}^A \wedge d\Omega^B + \frac{1}{2} S_{AB}\, (\tilde{E}^A \wedge \tilde{E}^B - \nu\, \tilde{F}^A \wedge \tilde{F}^B) \\ \nu = \begin{cases} 0 \ , & \text{Galilean case} \\ 1 \ , & \text{Lorentz case} \end{cases} \\ M = m + \nu \frac{1}{2} I_{AB}\, \Omega^A \Omega^B \quad . \end{cases}$$

§4.6 Canonical Distribution on $G(V)$; Motion in Space-Time.

(a) We defined in the last section the canonical system $(TG(V),\omega)$ and now proceed as outlined in §3.3 to calculate the distribution $p \to \ker \omega(p)$.

Let $Z \in T_p TG(V)$ be any vector; say,

(1) $$Z = \nu^a \underset{\sim}{e}_a + P^A \underset{\sim}{E}_A + Q^A \underset{\sim}{E}_A + \xi^a \frac{\partial}{\partial v^a} + \Phi^A \frac{\partial}{\partial \Omega^A} + \Psi^A \frac{\partial}{\partial U^A}$$

where the coordinates on $G(V)$ are (x^α, e^α_a) and those on $TG(V) = G(V) \times R^4 \times g$ are $(x^\alpha, e^\alpha_a; v^a, \Omega^A, U^A)$. By definition, $\ker \omega(p) = \{Z \in T_p TG(V) : Z \lrcorner\, \omega(p) = 0\}$. Now, in terms of (1),

(2) $$\begin{aligned} Z \lrcorner\, \omega = {} & (S_A \Phi^A - \tfrac{1}{2} S^{MN} R_{MNOr}\, \nu^r)\theta^o \\ & -(MQ_A + \tfrac{1}{2} S^{MN}\, R_{MNAr}\, \nu^r)\theta^A \\ & -(I_{AB}\Phi^B + S_{AB}P^B)\tilde{E}^A + (M\nu_A + \nu\, S_{AB}Q^B)\tilde{F}^A \\ & + (I_{AB}P^B - \nu^o\, S_A)\, d\Omega^A \quad , \end{aligned}$$

where we have made use of the fact that $\{\underset{\sim}{e}_a, \underset{\sim}{E}_A, \underset{\sim}{F}_A ; \frac{\partial}{\partial v^a}, \frac{\partial}{\partial \Omega^A}, \frac{\partial}{\partial U^A}\}$ and $\{\theta^a, \tilde{E}^A, \tilde{F}^A; dv^a, d\Omega^A, dU^A\}$ are dual orthonormal bases of $T_p TG(V)$ and $T^*_p TG(V)$ respectively, at each $p \in TG(V)$.

Therefore, $Z \in \ker \omega(p)$ iff $Z \lrcorner\, \omega(p) = 0$, iff

(3) $$P^A = \nu^o\, \Omega^A$$

(4) $$M\nu^A = -\nu\, S^A{}_B Q^B$$

(5) $$\Phi^A = -\nu^o\, \tilde{I}^{AB} S_{BC}\, \Omega^C$$
$$\tilde{I}^{AB} = (I^{-1})^{AB}$$

(6) $$MQ^A = -\frac{1}{2} S^{MN} R_{MNAr} \nu^r$$

and

(7) $$0 = S^{MN} R_{MNOr} \nu^{r}$$

where (5), implying

(8) $$S_A \Phi^A = 0 \quad ,$$

is used to calculate (7). Note also that (4) implies

(9) $$S_A \nu^A = 0$$

Substitute (6) into (4) to obtain

(10) $$M\nu^A - \nu \frac{1}{2} S^{MN} R_{MNrB} \nu^r \; S^{BA} = 0 \quad ,$$

an equation for ν^A in terms of the coordinates and ν^o only. In the Lorentz case, the solution of (10) is facilitated by first transforming to a frame such that $S^{AB} = s \;\; \delta^{AB}_{23}$, solving for ν^1 , ν^2 , ν^3 , and applying the inverse transformation. Equation (9) implies that ν^A has no component along S^A . We find then, that

(11) $$\nu^A = \frac{1}{2} M^{-2} \Delta^{-1} \nu^o \; S^{MN} R_{MNOB} S^{BA}$$

with

(12) $$\Delta = 1 + \frac{1}{4} M^{-2} \; S^{MN} R_{MNAB} S^{AB} \quad .$$

The assumption that Δ not being zero places a restriction (in the Lorentz case) on the magnitude of the spin ([27]). Physically, the test particle concept breaks down if $\Delta = 0$ (for fixed non-zero curvature),

so that the condition $\Delta \neq 0$ is not significant. We will later consider (12) as a guide to approximation in Schwarzschild geometry.

A vector $Z \in T_p TG(V)$, whose components (1) are given by (3), (5), (10) and (6) with ν^o, ξ^a, Ψ^A remaining arbitrary, lies in $\ker \omega(p)$; whence $\dim(\ker \omega(p)) = 8$ and $\dim(T_p TG(V)/\ker \omega(p)) = 12$. This holds at each $p = (x, e_a; v, \Omega, U) \in TG(V)$ (subject to (12); we interpret the evolution space now to be the open submanifold of $TG(V)$ defined by $\Delta \neq 0$), so that $\ker \omega$ is a foliation and the space of motions $M = TG(V)/\ker \omega$ is (assumed to be) a 12-dimensional symplectic manifold corresponding to the traditional 12-dimensional phase space of the rigid body.

(b) Motion in space-time. The motions of the system are the maximal integral submanifolds of the foliation $p \to \ker \omega(p)$. The equations describing the translational motion of the centre of mass of the body and the rotational motion are the equations of the projections of these integral submanifolds into space-time V , and $G(V)$. We begin with the following definitions: Let

(13) $$\lambda \to p(\lambda) = (x^\alpha(\lambda), e^\alpha_a(\lambda); v^a(\lambda), \Omega^A(\lambda), U^A(\lambda)) \in TG(V)$$

be an integral curve of a vector field Z in $\ker \omega$. By definition, $\lambda \to p(\lambda)$ satisfies

(14) $$\frac{d}{d\lambda} p(\lambda) = Z \lrcorner\, dp(\lambda) .$$

We use the notations

$$(15)\qquad \begin{cases} v^\alpha = \dfrac{dx^\alpha}{d\lambda} \\ \dot{} = v^\alpha \nabla_\alpha \end{cases} \quad \text{along} \quad \lambda \to x^\alpha(\lambda) \in V \quad .$$

The following space-time components of the corresponding body-frame quantities will be needed:

$$(16)\qquad \Omega^\alpha = e^\alpha_A \, \Omega^A$$

$$(17)\qquad S^\alpha = e^\alpha_A \, S^A$$

$$(18)\qquad S^{\alpha\beta} = e^\alpha_A \, e^\beta_B \, \epsilon^{AB}{}_C S^C = \epsilon^{\alpha\beta}{}_{\gamma\delta} \, e^\gamma_o \, S^\delta$$

$$(19)\qquad \Omega_{\alpha\beta} = \theta^A_\alpha \, \theta^B_\beta \, \epsilon_{ABC}\Omega^C = \epsilon_{\alpha\beta\gamma\delta} \, e^\gamma_o \, \Omega^\delta \quad .$$

We are interested only in the worldline of the centre of mass, $\lambda \to x^\alpha(\lambda)$, the corresponding development of the body frame, $\lambda \to e^\alpha_a(\lambda)$, the angular velocity, $\lambda \to \Omega^\alpha(\lambda)$, and the spin, $\lambda \to S^\alpha(\lambda)$. Explicitly, then (13) is:

$$(20)\qquad v^\alpha = \frac{dx^\alpha}{d\lambda} = Z \lrcorner \, dx^\alpha = Z \lrcorner \, (e^\alpha_a \theta^a) = \nu^a \, e^\alpha_a \quad .$$

Substitute from (11) to get

$$(21)\qquad v^\alpha = \nu^o (e^\alpha_o + \nu \, \tfrac{1}{2} \, M^{-2} \Delta^{-1} S^{\mu\nu} R_{\mu\nu\gamma\beta} \, e^\gamma_o \, S^{\beta\alpha}) \quad .$$

Next (see 4.3(16) for de^α_a),

$$\frac{de^\alpha_o}{d\lambda} = Z \lrcorner \, de^\alpha_o$$

so that

(22) $$\dot{e}^{\alpha}_{o} = e^{\alpha}_{A}\, Q^{A} \quad .$$

But, from (20), $\nu^{a} = \theta^{a}_{\alpha}\, v^{\alpha}$; using this in (6) we get

(23) $$M\dot{e}^{\alpha}_{o} = \frac{1}{2}\, S_{\mu}{}^{\nu} R^{\mu}{}_{\nu\beta}{}^{\alpha}\, v^{\beta} \quad .$$

(Note that $S^{MN} R_{MNr}{}^{A} \nu^{r}\, e^{\alpha}_{A} = S^{MN} R_{MNr}{}^{a} \nu^{r} e^{\alpha}_{a}$, in view of (7).) Similarly,

(24) $$\frac{de^{\alpha}}{d\lambda^{A}} = Z \lrcorner\, de^{\alpha}_{A} \quad .$$

Therefore

(25) $$\dot{e}^{\alpha}_{A} = \nu\, e^{\alpha}_{o}\, Q_{A} - \epsilon_{AB}{}^{C} P^{B} e^{\alpha}_{C} \quad .$$

Note that (in the Lorentz case) $e^{\alpha}_{A}\, e_{\alpha o} = e^{\alpha}_{A} g_{\alpha\beta} e^{\beta}_{o} = 0$; thus, $\dot{e}^{\alpha}_{A}\, e_{\alpha o} = -e^{\alpha}_{A}\, \dot{e}_{\alpha o}$ and (22) implies $\dot{e}^{\alpha}_{A}\, e_{\alpha o} = Q_{A}$. Substituting this value for Q_{A} into (25), and making use of (3) and (19), we find

(26) $$(\delta^{\alpha}_{\beta} + \nu\, e^{\alpha}_{o} e_{\beta o})\dot{e}^{\beta}_{A} = -\nu^{o}\, \Omega^{\alpha}{}_{\beta}\, e^{\beta}_{A} \quad .$$

However, noting that $\psi_{\alpha}\, e^{\alpha}_{o} = 1$ and $\psi_{\alpha}\Omega^{\alpha\beta} = \psi_{\alpha}\, e^{\alpha}_{A} e^{\beta}_{B} \Omega^{AB} = 0$ in the Galilei case, (26) can be equally well written as

(27) $$\bar{\delta}^{\alpha}_{\beta}\, e^{\beta}_{A} = -\nu^{o}\, \Omega^{\alpha}{}_{\beta}\, e^{\beta}_{A} \quad ,$$

where

(28) $$\bar{\delta}^{\alpha}_{\beta} = \begin{cases} \delta^{\alpha}_{\beta} - e^{\alpha}_{o}\, \psi_{\beta} \;, & \text{Galilei} \\ \delta^{\alpha}_{\beta} + e^{\alpha}_{o}\, e_{\beta o} \;, & \text{Lorentz} \end{cases}$$

is the projection operator onto the instantaneous rest space orthogonal to e_o .

Next, consider

$$\frac{d\Omega^\alpha}{d\lambda} = \frac{d}{d\lambda}\,(e^\alpha_A\Omega^A) = \Omega^A Z \lrcorner\, de^\alpha_A + e^\alpha_A\, Z \lrcorner\, d\Omega^A \quad .$$

Substituting (24) into this equation and using

$$\text{(29)} \qquad Z \lrcorner\, d\Omega^A = \Phi^A = -\nu^o\, \tilde{I}^{AB} S_{BC}\, \Omega^C \quad ,$$

we find that

$$\dot{\Omega}^\alpha = \nu\, e^\alpha_o\, Q_A\Omega^A - \nu^o\, e^\alpha_A\, \tilde{I}^{AB} S_{BC}\, \Omega^C \quad ;$$

whence, recalling (22) and (28),

$$\text{(30)} \qquad \tilde{\delta}^\alpha_\beta\, \dot{\Omega}^\beta = -\,\nu^o\, \tilde{I}^{\alpha\beta} S_{\beta\gamma}\, \Omega^\gamma \quad ,$$

where $\tilde{I}^{\alpha\beta} = e^\alpha_A e^\beta_B\, \tilde{I}^{AB}$; (30) is the space-time form of Euler's equations.

The last quantity we wish to consider is the spin S^α :

$$\frac{dS^\alpha}{d\lambda} = \frac{d}{d\lambda}\,(e^\alpha_A S^A) = S^A Z \lrcorner\, de^\alpha_A + e^\alpha_A\, I^A{}_B\, Z \lrcorner\, d\Omega^B \quad ,$$

which becomes, once the appropriate quantities are substituted,

$$\dot{S}^\alpha = \nu\, e^\alpha_o\, S_A\, Q^A$$

or using (28),

(31) $$\bar{\delta}^{\alpha}_{\beta}\, S^{\beta} = 0 \quad .$$

This essentially completes the description of the motion of a rigid test body in space-time. We close with some remarks and an example.

(c) Remarks.

(i) Due to the arbitrariness of the parameterization $\lambda \to p(\lambda)$ (13), ν^{o} is arbitrary; if λ is chosen to be proper time in the Lorentz case, or $\lambda = t = x^{o}$ (absolute time) in the Galilei case, ν^{o} is determined by (21).

(ii) In the Lorentz case, the condition $S_{\alpha}\, e^{\alpha}_{o} = 0$ or equivalently, $S^{\alpha\beta}\, e_{\beta o} = 0$, is one of the so-called spin supplementary conditions, and is automatically satisfied here because of the definition (17) and the fact that (e^{α}_{a}) is a Lorentz transformation (so that $e^{\alpha}_{o} g_{\alpha\beta} e^{\beta}_{A} = 0$).

From the beginning, these supplementary conditions were recognised as being necessary to give a complete set of equations of motion, along with those for linear velocity and spin. (For a historical critique, see Dixon [12]). Specifying such a condition on the spin has the effect of determining the centre of mass of the body, which, in relativity, is a rather tenuous concept. A comparison of two conditions, $S^{\alpha o} = 0$ and $S^{\alpha\beta}\, v_{\beta} = 0$, made in a practical setting in [4], shows that there may be a significant difference in the resulting equations of motion.

(iii) The quantities $M = m + \nu \frac{1}{2} S_{A}\, \Omega^{A}$ and $s = (S_{A} S^{A})^{1/2} = (S_{\alpha} S^{\alpha})^{1/2}$ are conserved along the solution curve. This follows by covar-

iantly differentiating M and S , and using the equations of motion, (30) and (31).

(d) A simple example in the Galilei case can be obtained by considering a flat Galilei structure with the induced metric $\gamma_{\Gamma\Lambda} \equiv \delta_{\Gamma\Lambda}$. The condition that a connection Γ on Gal(V) be Newtonian, $d\kappa = 0$ (cf. 4.3 (13), is equivalent to the local existence of a 1-form $A = A_\alpha \, dx^\alpha$ such that $\kappa = -\frac{1}{2} dA$, or $\kappa_{\alpha\beta} = \partial_{[\alpha} A_{\beta]}$. Define $B^\Gamma = \epsilon^{\Gamma\Lambda\Delta} \partial_\Lambda A_\Delta$ and $E^\Gamma = -\partial^\Gamma A_o$. Referring to 4.2 (10), we see that the nonvanishing components of the connection are $\Gamma^\Gamma_{oo} = -E^\Gamma$ and $\Gamma^\Gamma_o = -\frac{1}{2} \epsilon^\Gamma{}_{\Lambda\Delta} B^\Delta$. The curvature components can be written as

$$\text{(32)} \qquad R^\Gamma{}_{\sigma\lambda\mu} = 2\delta^o_\sigma (\partial_{[\lambda} \Gamma^\Gamma_{\mu]o} + \Gamma_{o\Sigma} \, \delta^o_{[\lambda} \Gamma^\Sigma_{\mu]o}) + \delta^\Delta_\sigma \, \delta^o_{[\mu} \partial_{\lambda]} \, \Gamma^\Gamma_{o\Delta} \quad .$$

Equation (20) becomes $v^o = \frac{dx^o}{d\lambda} = \nu^o$ which we choose to be 1, so that $x^o = \lambda = t$. From (20) and (23) we obtain $M\dot{v}^\Gamma = S_\Lambda \partial^\Gamma B^\Lambda$; in more detail,

$$\text{(33)} \qquad \begin{cases} M \dfrac{dv^\Gamma}{dt} = E^\Gamma + e^\Gamma{}_{\Lambda\Delta} \, v^\Lambda B^\Delta + S_\Lambda \partial^\Gamma B^\Lambda \\[2ex] M \dfrac{d\vec{v}}{dt} = \vec{E} + \vec{v}\times\vec{B} + (\nabla\vec{B})\cdot\vec{S} \quad . \end{cases}$$

Equation (31) implies that

$$\text{(34)} \qquad \begin{cases} \dfrac{dS^\Gamma}{dt} = \Gamma^\Gamma_{\Lambda\sigma} \, S^\Lambda v^\sigma = -\epsilon^\Gamma{}_{\Lambda\Delta} \, S^\Lambda B^\Delta \\[2ex] \dfrac{d\vec{S}}{dt} = \vec{B}\times\vec{S} \quad . \end{cases}$$

(e) Fermi transport of spin. Now that the equations for the centre of mass worldline and spin are known it is useful to see how (31), the spin-precession equation, can be interpreted. First, we recall the definitions of Fermi-Walker derivative and transport. Let $u = u^\alpha \partial_\alpha$ be the unit tangent vector to a timelike curve $\lambda \to x^\alpha(\lambda)$ in V and let $\xi = \xi^\alpha \partial_\alpha$ be any vector field. Then the <u>Fermi-Walker derivative</u> of ξ is

$$(35) \qquad (\delta_u \xi)^\alpha \equiv \begin{cases} \dot{\xi}^\alpha - \psi_\beta \xi^\beta \dot{u}^\alpha , & \text{Gal.} \\ \dot{\xi}^\alpha + u_\beta \xi^\beta \dot{u}^\alpha - \dot{u}_\beta \xi^\beta u^\alpha , & \text{Lor.} \end{cases}$$

ξ is said to be <u>Fermi-Walker transported</u> along the curve iff $\delta_u \xi = 0$ on the curve. Fermi-Walker transport preserves scalar products, so that an orthonormal frame at one point remains orthonormal when Fermi-Walker transported along a curve - such a frame-field is called a <u>Fermi frame</u>. A Fermi frame is the natural space-time analogue of a non-rotating frame. If (e_a) is a Fermi frame with $e_0 = u$, then (e_A) is a non-rotating triad in the instantaneous rest space orthogonal to u.

Applying (35) to the spin vector S^α we get

$$(36) \qquad (\delta_u S)^\alpha = \begin{cases} 0 , & \text{Gal.} \\ \dot{S}^\alpha + u_\beta S^\beta \dot{u}^\alpha - \dot{u}_\beta S^\beta u^\alpha , & \text{Lor.} \end{cases}$$

where (31) was used for the Galilei case. We obtain immediately the result that spin is Fermi-Walker transported in the Galilei case. To compute (36) for the Lorentz case it is necessary to first obtain e_0^α in terms of u^α in order to make use of (31). This is done by inverting equation (21).

All considerations now refer to the Lorentz situation only.

Define

(37) $$\tilde{u}^{\alpha} \equiv \frac{v^{\alpha}}{v^{0}} \quad , \text{ and}$$

(38) $$\kappa \equiv \frac{1}{2}\, m^{-2} \Delta^{-1} \quad .$$

The solution of (21) for e^{α}_{0} in terms of $\tilde{u}^{\alpha}$ can be obtained by first decomposing e^{α}_{0} as

(39) $$e^{\alpha}_{0} = \tilde{u}^{\alpha} + g^{\alpha}, \quad g_{\alpha}\tilde{u}^{\alpha} = 0.$$

Substitution of (39) into (21) yields

(40) $$g^{\alpha} + \kappa\, S^{\mu\nu} R_{\mu\nu\rho\sigma}\, \tilde{u}^{\rho} S^{\sigma\alpha} + \kappa\, S^{\mu\nu} R_{\mu\nu\rho\sigma} g^{\rho}\, S^{\sigma\alpha} = 0$$

from which it can be derived that

(41) $$g^{0} \equiv \theta^{0}_{\alpha} g^{\alpha} = 0.$$

Thus, in frame components (40) becomes a 3-dimensional problem:

(42) $$g^{A} + \kappa\, S^{MN} R_{MNkL} g^{k} S^{LA} + \kappa\, S^{MN} R_{MNkL}\, \tilde{u}^{k} S^{LA} = 0$$

which can be solved using the trick employed to derive (21):

(43) $$g^{A} = -\frac{1}{2}\, M^{-2}\, S^{MN} R_{MNrB}\, \tilde{u}^{r}\, S^{BA}.$$

Returning to (39), we see that

(44) $$e^{\alpha}_{0} = \tilde{u}^{\alpha} - \frac{1}{2}\, M^{-2}\, S^{\mu\nu} R_{\mu\nu\rho\sigma} \tilde{u}^{\rho}\, S^{\sigma\alpha} \quad .$$

Now, define

(45) $$\lambda^{2} = \frac{1}{4}\, M^{-4}\, S^{\mu\nu} R_{\mu\nu\rho\sigma} u^{\rho} S^{\sigma}_{\alpha}\, S^{\alpha\beta} R_{\beta\gamma\delta\varepsilon} u^{\nu}\, S^{\delta\varepsilon} \quad .$$

Then using the definition (37) of $\tilde{u}$ together with the result (44),

$$\frac{(v^0)^2}{v_\alpha v^\alpha} = 1 - \lambda^2, \tag{46}$$

from which it follows

$$(1 - \lambda^2)^{\frac{1}{2}} e_0^\alpha = u^\alpha - \tfrac{1}{2} M^{-2} S^{\mu\nu} R_{\mu\nu\rho\sigma} u^\rho S^{\sigma\alpha} \tag{47}$$

(We shall see that in reasonable situations λ will be small compared to c = speed of light = 1.)

Returning to the case in hand, note that (21) and $S^\alpha e_{\alpha 0} = 0$ imply

$$S^\alpha u_\alpha = 0, \tag{48}$$

so that

$$\dot{S}^\alpha u_\alpha = -S^\alpha \dot{u}_\alpha . \tag{49}$$

Equations (36), (48) and (49) show that

$$(\delta_u S)^\alpha = (\delta^\alpha_\beta + u^\alpha u_\beta)\dot{S}^\beta. \tag{50}$$

For brevity, introduce

$$R_\alpha \equiv \tfrac{1}{2} M^{-2} S^{\mu\nu} R_{\mu\nu\rho\alpha} u^\rho, \tag{51}$$

so that (31) becomes, using (47) and (50),

$$\begin{aligned}(\delta_u S)^\alpha = {} & R_\rho (u^\alpha S^\rho_\beta + u_\beta S^{\rho\alpha})\dot{S}^\beta + \lambda^2 u^\alpha u_\beta \dot{S}^\beta \\ & - \lambda^2 R_\rho (u^\alpha S^\rho_\beta + u_\beta S^{\rho\alpha})\dot{S}^\beta .\end{aligned} \tag{52}$$

Thus, in the Lorentz case, spin is not Fermi-transported along the centre

of mass worldline. In §5.9 we will see in a typical situation how much $\delta_u S$ differs from zero.

CHAPTER V

Constrained Motion of a Rigid Body

§5.1 Introduction.

(a) Before proceeding with the main subject matter of this chapter, we make a few comments of a general nature. The usual discussions of constraints in mechanics often don't make quite clear what issues are involved and how. We believe that a discussion based on a geometrical formulation of Lagrangian dynamics can provide both a natural conceptual framework as well as the necessary apparatus for dealing with concrete problems.

Standard usage of the verb "constain" implies that something is being constrained to something else; otherwise said "constrain" is an operator. In our context, to a given dynamical system, this operator assigns a new dynamical system in a manner which depends on the particular description of such systems.

For example, the evolution space of the traditional Lagrangian mechanics is $R \times TQ$, Q being the n - dimensional configuration space of some dynamical system. An associated constrained dynamical system is then understood to be obtained from a sub-fibre bundle N over R of the fibre bundle $R \times TQ \to R$: For each $t \varepsilon R$ we assume that there is given an $(m \leq 2n)$ - dimensional submanifold N_t of TQ so that N is obtained as $N = \bigcup_t N_t$. Reducing the evolution space from $R \times TQ$ to N will have an effect on the dynamics, in that constraint forces, a priori unknown in general, will now appear in the equations of

motion. These constraint forces are often thought of as being the agency which restricts the motion to N. If N is a product, $N = R\times M$ where M is a sub-bundle of TQ, the constraint is said to be <u>scleronomous</u> (time-independent); otherwise, it is <u>rheonomous</u> (time-dependent). Furthermore, if each $N_t \subset TQ$ is derivable from constraints in Q, $N_t = TC_t$, $C_t \subset Q$ being a submanifold of Q, then the constraint N is called <u>holonomic</u>; otherwise, it is called non-holonomic. The generalization of this example to a homogeneous description of mechanics, the evolution space being $T(R\times Q)$, is straightforward, and we consider it to be done with the understanding that Q now contains the time coordinate. The constraint manifold N is then a submanifold of TQ.

The case that $\dim N < \dim TQ$ (holonomic or velocity-dependent constraints) is often described in terms of a smooth map $\phi : TQ \to R^m$ where $m < n = \dim Q$; that is, $N = \phi^{-1}(\text{const.})$. If $\phi_{*}|_p : T_pTQ \to T_{\phi(p)}R^m$ is surjective for all $p \in N$, then N is an embedded submanifold of TQ and has dimension $2n-m$ ([59]).

When N is holomonic so that $N = TM$ for some submanifold $M \subset Q$, describe this M (since all questions studied here are local we will denote again by M a certain open submanifold $U \subset M$) by a submersion $\phi : Q \to R^m$ $(m < n)$, $M = \phi^{-1}(0)$, say. A vector $v \in T_qQ$ is in T_qM iff

$$\bar{\phi}^{\Gamma}(q,v) \equiv v \lrcorner\, d\phi^{\Gamma}(q) = 0$$

where ϕ^{Γ} is the Γ-th component function of ϕ. Thus we can describe N by $2m$ functions:

$$N = \{(q,v) \in TQ : \phi^{\Gamma}(q) = 0 = \bar{\phi}^{\Gamma}(q,v)\}.$$

It is clear that the map $\bar{\phi} \equiv (\phi^{\Gamma}, \bar{\phi}^{\Gamma}) : TQ \to R^m$ is a submersion whenever ϕ is (compute the Jacobian).

(b) Historical development of constraints. An important category of constraints are those arising from systems described by degenerate Lagrangians. We now try to indicate the difficulties involved and what they lead to.

The Hamiltonian formalism of mechanics appears to be the most convenient as a point of departure in the quantization of classical systems. As well as being the traditional way of proceeding from classical to quantum mechanics, the Hamiltonian formalism gives us a canonically defined Lie algebra of observables (cf. §4.4d) which is the basis of some modern prescriptions for quantization (for example, Hermann [21]). On the other hand, while the requirement that a theory be relativistic is easy enough to formulate in terms of a Lagrangian and its associated action principle, this is usually not the case in a Hamiltonian formulation (this becomes especially apparent in relativistic field theory where the Hamiltonian formalism treats the time coordinate and space coordinates very differently ([11]; [42] page 287)).

The procedure outlined in §2.5 which constructs a canonical Lagrangian structure on TQ depended on the Lagrangian L being regular (2.5 (13)) to pull back the canonical symplectic form ω_o on T^*Q to TQ. In [1] it is shown that if the fibre derivative FL is a diffeomorphism, then the Hamiltonian and Lagrangian formulations of

mechanics are equivalent - the projections of the solution curves, from T^*Q and TQ respectively, into Q coincide. In this case, we can readily combine the advantages of the possiblities of relativistic covariance and canonical quantization by taking the Lagrangian formulation as basic.

Unfortunately, there are very few physical systems of interest for which L is regular.[10)] This, together with the need to quantization led a number of people into considering methods for dealing with such degenerate Lagrangian systems ([10], [11], [26], [31], [45], [46]).

(c) The purpose of this chapter is to investigate the constrained motion of a rigid test body in curved (Newtonian or Lorentzian) space-time. The next section shows how to geometrically view the traditional formulation of Lagrangian constraints and leads to an inhomogeneous formulation of symplectic constraints. §5.3 outlines a homogeneous formulation of symplectic constraints while §5.4 applies this to the rigid body problem; unfortunately, the relativistic equations of motion are not completely determined in this case.

The reformulation of the problem as a second order equation in §5.5 in the spirit of [29] (Appendix) and [1] not only provides us with a deterministic set of equations of motion, but also a plausible reason for the failure in §5.4. §5.7 examines some simple examples (Galilei space-time, Minkowski space and a DeSitter universe) of motion constrained to a worldline. Finally, after showing how to obtain the general solution, we give a "practical" example: motion constrained to circular orbits in Schwarzschild geometry.

§5.2 Geometrization of the Lagrange Multiplier Formalism.

(a) In this section we assume all constraints to be holonomic. Although holonomic constraints can be eliminated right away by making coordinate transformations in the state space, it is instructive to use the geometrical interpretation of the traditional Lagrange procedure to do this.

(b) Let Q denote the configuration space of a dynamical system with $\dim Q = n$. Following the notation of §2.5, TQ is the state space and $L : TQ \to R$ the Lagrangian (assumed to be regular). The motion of the system is assumed to be subject to the constraint defined by $f : Q \to R^m$ (f is assumed to be a submersion):

(1) $$f^A(q) = 0 \quad , \quad A = 1,\ldots,m < n.$$

A motion of the system is represented by a curve $R \to Q$, $\lambda \to x^\alpha(\lambda)$ such that its lift into TQ , $\lambda \to (q^\alpha(\lambda) \, , \frac{dq^\alpha}{d\lambda}(\lambda))$, satisfies the Euler-Lagrange equations:

(2) $$\frac{d}{d\lambda}\left(\frac{\partial L}{\partial v^\alpha}\right) - \frac{\partial L}{\partial q^\alpha} = \lambda_A \frac{\partial f^A}{\partial q^\alpha}$$

(3) $$v^\alpha(\lambda) = \frac{dq^\alpha}{d\lambda}$$

where the λ_A are the Lagrange multipliers, functions on TQ . Equations (2) and (3) are subject to the constraints (1), as well as the compatibility conditions

(4) $$\frac{df^A}{d\lambda} = 0 \quad .$$

These are essentially equivalent, in the holonomic case, to (1), but we find it convenient to state them both here.

Now, being a lift, $\lambda \to (q^\alpha(\lambda), v^\alpha(\lambda))$ is an integral curve of a vector field Z on TQ such that

$$Z \lrcorner \, dq^\alpha = v^\alpha \quad . \tag{5}$$

If $F : TQ \to R$ is any function of (q,v) then

$$Z \lrcorner \, dF = Z(F) = \frac{dF}{d\lambda} \tag{6}$$

by definition of $\lambda \to (q^\alpha, v^\alpha)$ being an integral curve of Z. Using (6), (2) becomes

$$Z\left(\frac{\partial L}{\partial v^\alpha}\right) = \frac{\partial L}{\partial q^\alpha} + \lambda_A \frac{\partial f^A}{\partial q^\alpha} \quad , \tag{7}$$

so that

$$Z\left(\frac{\partial L}{\partial v^\alpha}\right) dq^\alpha = \frac{\partial L}{\partial q^\alpha} dq^\alpha + \lambda_A \, df^A \quad , \tag{8}$$

since $df^A = \frac{\partial f}{\partial q^\alpha} dq^\alpha$. Define the 1-form $\theta \in \mathcal{X}^*(TQ)$ by

$$\theta = \frac{\partial L}{\partial v^\alpha} dq^\alpha \tag{9}$$

(cf: §2.5) and also note that

$$L_Z(\theta) = Z\left(\frac{\partial L}{\partial v^\alpha}\right) dq^\alpha + L_Z(dq^\alpha) \frac{\partial L}{\partial v^\alpha} \tag{10}$$

where L_Z denotes the Lie derivative. Use the definition (9) and the formula (10) in (8) to get:

(11) $$L_Z(\theta) - L_Z(dq^\alpha)\frac{\partial L}{\partial v^\alpha} = \frac{\partial L}{\partial q^\alpha} dq^\alpha + \lambda_A\, df^A \quad .$$

We recall a standard formula: if α is a k-form,

(12) $$L_Z(\alpha) = Z \lrcorner\, d\alpha + d(Z \lrcorner \alpha) \quad .$$

Applied to (5),

(13) $$L_Z(dq^\alpha) = Z \lrcorner\, d(dq^\alpha) + d(Z \lrcorner dq^\alpha) = dv^\alpha \quad .$$

The 2-form ω is defined by

(14) $$\omega = -d\theta \quad ;$$

hence,

(15) $$L_Z(\theta) = -Z \lrcorner\, \omega + d(v^\alpha \frac{\partial L}{\partial v^\alpha})$$

using (12) and (5). Substituting (13) and (15) into (11),

$$-Z \lrcorner\, \omega + d(v^\alpha \frac{\partial L}{\partial v^\alpha}) - \frac{\partial L}{\partial v^\alpha} dv^\alpha - \frac{\partial L}{\partial q^\alpha} dq^\alpha = \lambda_A\, df^A \quad ,$$

or

(16) $$Z \lrcorner\, \omega - dE = \lambda_A\, df^A$$

where

(17) $$E = v^\alpha \frac{\partial L}{\partial v^\alpha} - L$$

is the energy (cf. §2.5).

If we denote by M the constraint manifold in Q ,

$$M = \{q : f^A(Q) = 0\} \quad , \tag{18}$$

the "lifted" constraint manifold N in TQ is just

$$N = TM = \{(q,v) \in TQ : f^A(q) = 0 \;,\; \bar{f}^A(q,v) = 0\} \quad , \tag{19}$$

where we have defined the map $\bar{f} : TQ \to R^m$ by

$$\begin{cases} \bar{f}^A(q,v) = v \lrcorner \, df^A(q) \quad , \quad \text{or} \\ \\ \bar{f}^A = v^\alpha \dfrac{\partial}{\partial x^\alpha} f^A \quad . \end{cases} \tag{20}$$

The condition that the motion stays in N is that Z be tangent to N and is equivalent to

$$\begin{cases} Z \lrcorner \, df^A = 0 \\ \\ Z \lrcorner \, d\bar{f}^A = 0 \quad . \end{cases} \tag{21}$$

Let $\iota : N \to TQ$ be the embedding of N as a submanifold. If α is a k-form on TQ then

$$\iota^*(d\alpha) = d(\iota^*\alpha) \quad . \tag{22}$$

We can state our results on N intrinsically as follows: From (16) and using (22),

$$\tilde{Z} \lrcorner (\iota^* \omega) - d(\iota^* E) = \lambda_A d(\iota^* f^A) = 0$$

since $\iota^* f^A = f^A\big|_N = 0$, and where $\iota_*(\tilde{Z}) = Z\big|_N$ (Z is tangent to N). Hence,

$$\tilde{Z} \lrcorner \tilde{\omega} - d\tilde{E} = 0 \quad \text{on} \quad N \tag{23}$$

where $\tilde{\omega} = \iota^* \omega$. (cf. §2.5 (19).)

(c) Now, let us suppose that we are given the system (23) on N defined by (19). We try to reverse the process leading to (23). Basic to this is the following fact: if $\tilde{\sigma} \in \mathcal{X}^*(N)$ is a 1-form on $N \subset TQ$ then $\tilde{\sigma} = 0$ is equivalent to $\tilde{\sigma} = \iota^* \sigma$ where $\sigma = \lambda_A \, df^A + \mu_A \, d\bar{f}^A$ on TQ with $\{\lambda_A, \mu_A\}$ arbitrary functions on TQ (this follows from (22) and (19)). Apply this to (23):

$$\begin{cases} Z \lrcorner \omega + dE = \lambda_A \, df^A + \mu_A \, d\bar{f}^A \\ f^A = 0 \quad , \quad \bar{f}^A = 0 \\ Z \lrcorner df^A = 0 \quad , \quad Z \lrcorner d\bar{f}^A = 0 \end{cases} \tag{24}$$

The steps (8) to (17) are reversible so that (24) is equivalent to

$$Z\left(\frac{\partial L}{\partial v^\alpha}\right) dq^\alpha = \frac{\partial L}{\partial q^\alpha} \, dq^\alpha + \lambda_A \, df^A + \mu_A \, d\bar{f}^A \quad ;$$

or, expanding df^A and $d\bar{f}^A$ in terms of dx^α and dv^α,

$$(25)\quad \begin{cases} Z\left(\dfrac{\partial L}{\partial v^\alpha}\right) = \dfrac{\partial L}{\partial q^\alpha} + \lambda_A \dfrac{\partial f^A}{\partial q^\alpha} + \mu_A v^\beta \dfrac{\partial^2 f^A}{\partial q^\beta \partial q^\alpha} \\[2em] 0 = \mu_A \dfrac{\partial f^A}{\partial q^\alpha} \quad . \end{cases}$$

By assumption, $\left(\frac{\partial f^A}{\partial q^\alpha}\right)$ has maximal rank $m < n$ so that $\mu_A \frac{\partial f^A}{\partial q^\alpha} = 0$ implies $\mu_A = 0$. Thus (25) is equivalent to $\frac{d}{d\lambda}\left(\frac{\partial L}{\partial v^\alpha}\right) - \frac{\partial L}{\partial q^\alpha} = \lambda_A \frac{\partial f^A}{\partial q^\alpha}$, which is (2).

(d) We can summarize the above discussion by the

<u>Theorem</u>: Let Q be a configuration space $\dim Q = n$, TQ state space and $L : TQ \to R$ a regular Lagrangian describing a physical system subject to the constraints $f^A = 0$, $A = 1,\ldots,m<n$ where $f : Q \to R^m$ is assumed to be nondegenerate on $M = \{q \in Q : f(q) = 0\}$. Let $N = TM$. The following two descriptions of constrained motion of the system are equivalent:

(i) the motion is a curve $\lambda \to (q^\alpha(\lambda), v^\alpha(\lambda))$ in TQ satisfying the Euler-Lagrange equations

$$\left.\begin{array}{ll} & \dfrac{d}{d\lambda}\left(\dfrac{\partial L}{\partial v^\alpha}\right) - \dfrac{\partial L}{\partial q^\alpha} = \lambda_A \dfrac{\partial f^A}{\partial q^\alpha} \\[1.5em] & v^\alpha(\lambda) = \dfrac{dq^\alpha}{d\lambda}(\lambda) \\[1.5em] \text{subject to} & f^A(q) = 0 \\[1.5em] \multicolumn{2}{l}{\text{where the } \lambda_A \text{ are functions on } TQ \text{ to be determined.}} \end{array}\right\}$$

The λ_A are called the generalized constraint forces.

(ii) the motion is an integral curve of a vector field $\tilde{Z}$ on N satisfying

$$\tilde{Z} \lrcorner \tilde{\omega} - d\tilde{E} = 0 \qquad \text{(on N)}$$

where $\tilde{\omega}$ and $\tilde{E}$ are associated to $\iota^* L$ as in §2.5.

§5.3 Geometrical Theory of Constraints.

From now on, the mechanical systems we will be dealing with will be described more generally. By this we mean that no longer will a Lagrangian describe the dynamics; but, we will assume that a 1-form θ such that $\omega = d\theta$ is a (pre-)symplectic form, or even a given presymplectic form ω, is the basis of dynamics. (Cf. §4.5.)

Constrained dynamical systems will first be described abstractly; then we specialize to the case that constraints are holonomic. It may have been noticed that in the last section the geometrical formalism induced by the Lagrangian could not describe nonholonomic constraints. A geometrical theory of mechanics which appears to be especially suited to handling nonholonomic constraints has been developed by Vershik and Faddeev [58]. Essentially they use a geometrical structure on TQ itself (where Q is a configuration space) consisting of a mixed-rank second-order tensor τ on TQ, a 2-form Ω_L associated to a Lagrangian L on TQ via τ, and a vector field X_L on TQ defined by means of Ω_L and L, whose integral curves represents the motion. (Klein [23] also considers this particular geometrization of mechanics in a more general setting - he recognizes the geometry as being an almost-tangent

structure). Constraints are then introduced as a submanifold $M \subset TQ$ and related to the mechanics by τ and the sub-bundle $TM \subset TTQ$.

Following §4.4d, we define a <u>dynamical system</u>[11)] to be a pair (P,ω) where P is an n - dimensional manifold, the evolution space of the system, and ω is a presymplectic form defined on P. A pair $((P,\omega), \iota : P' \to P)$ consisting of a dynamical system (P,ω) together with an embedded submanifold P', $\iota : P' \to P$ being the embedding, is called a <u>constrained dynamical system</u>. Unless explicitly stated otherwise, P' is assumed to be contained in P as a subset.

Given a constrained dynamical system $((P,\omega), \iota : P' \to P)$ we can define the following four vector spaces at each $p \in P$:

(i) T_pP : the tangent space to P at p

(ii) T_pP' : the tangent space to P' at p

(iii) $N_p = \{v \in T_pP : \iota^*(v \lrcorner \omega) = 0\}$

(iv) $K_p = N_p \cap T_pP' = \{v \in T_pP' : v \lrcorner (\iota^*\omega) = 0\} = \ker(\iota^*\omega(p))$.

Using the dimensions of these vector spaces (assuming $\dim P = n$ fixed) we can define two integer-valued functions on P:

$$p \to \dim K_p$$

$$p \to \dim N_p - \dim K_p \quad .$$

We assume that these functions are constant on P (an indication of what happens otherwise is found in [45] and [46]). We will denote by K both the sub-bundle $\bigcup_{p' \in P'} K_{p'}$, of TP' and also the distribution

$p' \to K_{p'}$. Since $\dim K$ is constant on P', $K = \ker(\iota^*\omega)$ is a completely integrable distribution and hence defines a foliation on P' (see §4.4d). Points of the quotient set P'/K represent motions (or histories) of the system.

It is convenient to formulate in P the condition that $v \in K_p$ for $p \in P'$. At each point $p \in P'$ let $P'^{\perp}_p$ denote the annihilator of $T_pP' \subset T_pP$,

(1) $$P'^{\perp}_p = \{\sigma \in T^*_pP : \sigma(v) = 0 \quad \text{for all} \quad v \in T_pP'\} \quad .$$

Then for each $p \in P'$ and for all $\sigma \in T^*_pP$,

(2) $$\iota^*(\sigma) = 0 \quad \text{iff} \quad \sigma \in P'^{\perp}_p \quad .$$

The equation (at $p \in P'$)

(3) $$v \lrcorner (\iota^*\omega) = 0$$

is equivalent to

$$0 = \iota^*((\iota_* v) \lrcorner \omega) = \iota^*(v \lrcorner \omega)$$

where we denote $\iota_* v$ again by v. But, by (2), this is precisely equivalent to saying that (3) is equivalent to

(4) $$\begin{cases} v \lrcorner \omega \in P'^{\perp}_p & \text{with} \\ \\ v \in T_pP' & . \end{cases}$$

More concretely, let Q be a n - dimensional (extended) configuration space of some system and let $P = TQ$. Suppose that $f : Q \to R^m$ $(m<n)$ is regular on $C = f^{-1}(0) = \{q \in Q : f(q) = 0\}$. Denote the component functions of f by $f^A : Q \to R$, $A = 1,\ldots,m$ and let $(q^\alpha ; \alpha = 0,\ldots,n-1)$ be a local coordinate system at some point in Q. Then $P' = TC$ is described as the set of zeros of the map $\bar{f} : P = TQ \to R^m \times R^m$, $\bar{f}(q,v) = (f^A(q) , \bar{f}^A(q,v) = v \lrcorner\, df^A(q))$; or locally, $\bar{f}^A = v^\alpha \dfrac{\partial f^A}{\partial q^\alpha}$. As in §5.2, $P'^\perp$ is spanned locally by $\{df^A, d\bar{f}^A\}$.

The conditions (4) can now be stated as

$$
(5) \qquad \begin{cases}
v \lrcorner\, \omega = \lambda_A \, df^A + \mu_A \, d\bar{f}^A \\[2ex]
v \lrcorner\, df^A = 0 \; , \; v \lrcorner\, d\bar{f}^A = 0 \\[2ex]
f^A = 0 \quad , \quad \bar{f}^A = 0 \\[2ex]
\lambda_A \text{ and } \mu_A \text{ to be determined}
\end{cases}
$$

§5.4 Application to Constrained Rigid Body Motion.

(a) The general discussion of the last section and in particular the equations 5.3(5) will now be specialized to the evolution space $P = G(V)$ and the presymplectic form discussed in §4.5 and §4.6. We are interested in seeing the effect of constraining a rigid test body at its centre of mass to a specified worldline and in particular a geodesic (with respect to the connection on $G(V)$ - see §4.5). It is expected

that no effect is seen in the Galilei case, but that a precession of the spin axis will occur in some (general) relativistic situations. Unfortunately, we don't see this now but have to wait until §5.8 and §5.9 since the present formalism leads to equations which do not completely determine the motion of the system.

(b) The definitions and conventions used here are those found in §4.5. A worldline,[12)] represented by $\lambda \to \psi^{\alpha}(\lambda)$ locally, in V can be locally described by $\{x \in V : \phi^{\Gamma}(x) = 0\}$ where for example

$$(1)\qquad \begin{cases} \phi^{\Gamma}(x^{\alpha}) = x^{\Gamma} - \psi^{\Gamma}(g(x^{o})) \\ \\ x^{o} = \psi^{o}(\lambda) \quad \text{iff} \quad \lambda = g(x^{o}) \quad . \end{cases}$$

We denote the lift $\pi^{*}\phi^{\Gamma}$ of ϕ^{Γ} into $G(V)$ again by ϕ^{Γ} (so, $\phi^{\Gamma}(x,e_a) = \phi^{\Gamma}(x)$).

The equations 5.3(5) are now written as:

$$(2)\qquad \phi^{\Gamma} = 0$$

$$(3)\qquad \bar{\phi}^{\Gamma} \equiv v^{r} \nabla_{r} \phi^{\Gamma} = 0$$

$$(4)\qquad Z \lrcorner\, d\phi = 0$$

$$(5)\qquad Z \lrcorner\, d\bar{\phi} = 0$$

$$(6)\qquad Z \lrcorner\, \omega = \lambda_{\Gamma}\, d\phi^{\Gamma} + \mu_{\Gamma}\, d\bar{\phi}^{\Gamma} \quad .$$

The vectorfield Z has the form §4.6(1):

$$Z = \nu^a \underset{\sim}{e}_a + P^A \underset{\sim}{E}_A + Q^A \underset{\sim}{F}_A + \xi^a \frac{\partial}{\partial v^a} + \Phi^A \frac{\partial}{\partial \Omega^A} + \Psi^A \frac{\partial}{\partial U^A} .$$

Introduce the notation:

(7) $$F^\Gamma_a \equiv \phi^\Gamma_{\ |a} = e^\alpha_a \frac{\partial \phi^\Gamma}{\partial x^\alpha}$$

(8) $$f^\Gamma \equiv F^\Gamma_o$$

(9) $$(G^A_\Gamma) \equiv (F^\Gamma_A)^{-1}$$

(10) $$g^A \equiv -G^A_\Gamma f^\Gamma .$$

We have assumed that ϕ is regular, hence (F^Γ_a) has maximal rank.

Denote by M the submanifold of $TG(V)$ defined by (2) and (3). Using the formulae §4.3(17), §4.3(18) we calculate the differentials of the constraints:

(11) $$d\phi^\Gamma = F^\Gamma_r \theta^r$$

(12) $$d\bar{\phi} = v^r F^\Gamma_{r|s} \theta^s - F^\Gamma_K v_L \epsilon^{KL}_{\ \ B} \tilde{E}^B$$

$$+ (v^o F^\Gamma_B + \nu f^\Gamma v_B)\tilde{F}^B + F^\Gamma_r dv^r .$$

Equations (3) and (4) are computed to be $F^\Gamma_a v^a = 0$ or $v^A = v^o g^A$, and $F^\Gamma_a \nu^a = 0$ or $\nu^A = \nu^o g^A$, respectively, and therefore imply

$$(13) \qquad \nu^A = \frac{\nu^o}{v^o} v^A \quad ,$$

while (5) works out to be

$$(14) \qquad F^\Gamma_r \xi^r + (F^\Gamma_B + \nu f^\Gamma v_B)Q^B - F^\Gamma{}_K v_L \epsilon^{KL}{}_B P^B$$

$$+ v^r F^\Gamma_{r|s} \nu^s = 0 \quad .$$

Finally, $Z \lrcorner \omega$ was computed in §4.6(2), so that the use of (11) and (12) enables us to express (6) as follows:

$$(15) \qquad S_A \Phi^A + \frac{1}{2} S^{MN} R_{MNro} \nu^r = \lambda_\Gamma f^\Gamma + \mu_\Gamma F^\Gamma_{a|o} v^a$$

$$(16) \qquad \frac{1}{2} S^{MN} R_{MNrA} \nu^r - MQ_A = \lambda_\Gamma F^\Gamma_A + \mu_\Gamma F^\Gamma_{a|A} v^a$$

$$(17) \qquad I_{AB}\Phi^B + S_{AB}P^B = \mu_\Gamma F^\Gamma_K v_L \epsilon^{KL}{}_A$$

$$(18) \qquad M\nu^A + \nu S_{AB}Q^B = \mu_\Gamma(v^o F^\Gamma_A + \nu f^\Gamma v_A)$$

$$(19) \qquad I_{AB}P^B - \nu^o S_{AB} = 0$$

$$(20) \qquad 0 = \mu_\Gamma F^\Gamma_a \quad .$$

In the Lorentz case (ν=1) (13) to (20) do not provide enough conditions to fully determine the motion (the difficulty being that the component of Q^A along S^A is not determinable; the equation for $\dot{e}^\alpha_o$ depends on Q^A, hence e^α_o cannot be determined). In the Galilei case (ν=0) we can obtain the equations of motion as follows:

Equation (20) implies that

(21) $$\mu_\Gamma = 0 \quad ,$$

since (F^Γ_a) has maximal rank. This implies that:

(22) $$\nu^A = 0$$

(23) $$P^A = \nu^o \, \Omega^A$$

(24) $$\dot{\Phi}^A = -\nu^o \, \tilde{I}^{AB} \, S_{BC} \, \Omega^C \quad .$$

Let λ be a parameter of an integral curve of Z (cf. §4.6b; in particular, 4.6(13) and 4.6(14)).

(25) $$v^\alpha = \frac{dx^\alpha}{d\lambda} = Z \lrcorner \, dx^\alpha = \nu^o \, e^\alpha_o$$

determines e^α_o along the a priori given constraint worldline $\lambda \to x^\alpha(\lambda)$. (Note that this determines Q^A : $\dot{e}^\alpha_o = e^\alpha_A \, Q^A$, giving Q^A in terms of (25).) The remainder of the equations of motion are obtained in exactly the same way as in §4.6b. In essence, constraining the centre of mass of a rigid body in Newtonian theory does not affect rotational motion of the body.

The discussion in the next section will both give us a workable model in the Lorentz case and explain why this case doesn't work here, whereas the Galilei does.

§5.5 Reformulation of the Problem as a Second-Order System.

(a) Many dynamical systems are describable as second-order equations (SOE) on some manifold M ; indeed, the classical Lagrange equations were formulated as such in §2.5. The remainder of our discussion of the rigid body is based on the consequences of assuming such a second-order principle. It turns out that the condition that the vector field Z found in §4.6a be a SOE imposes new constraints in the Lorentz case, but results in no new conditions in the Galilei case. We begin with a discussion of the SOE formulation outlined in the Appendix to [29].

(b) The setting now is the same as §4.5 and §4.6a. We assume that a motion of the system is determined by a second-order equation Z on $TG(V)$; that is, Z is of the form

$$Z = \mu(v^a \underset{\sim}{e}_a + \Omega^A \underset{\sim}{E}_A + U^A \underset{\sim}{F}_A) + \xi^r \frac{\partial}{\partial v^r} + \Phi^A \frac{\partial}{\partial \Omega^A} + \Psi^A \frac{\partial}{\partial U^A} \tag{1}$$

where μ , ξ^r , Φ^A , Ψ^A are arbitrary (at this stage). We still require that $Z \in \ker \omega$,

$$Z \lrcorner\, \omega = 0 \quad ; \tag{2}$$

but substitution of (1) into (2) and using the expression 4.5(14) for ω shows that (2) can be solved only on the submanifold $N \subset TG(V)$ defined as the set of zeros of the functions:

$$\Sigma^o = v^o - 1 \tag{3}$$

$$\Sigma^A = Mv^A + \nu\, S^A{}_B\, U^B \tag{4}$$

$$T^A = MU^A - R_r{}^A v^r \quad , \tag{5}$$

where we have introduced the shorthand notation

$$R_{ab} = \frac{1}{2} S^{MN} R_{MNab} \quad . \tag{6}$$

Differentation of (3) to (5) shows that Σ^a , T^A are independent on N . One additional condition arising from (2),

$$v^a R_{ao} = 0 \quad , \tag{7}$$

is a consequence of multiplying (5) by v_A and noting that (4) implies

$$v_A U^A = 0 \quad \text{on} \quad N. \tag{8}$$

Where $TG(V)$ is a 20-dimensional manifold, the new evolution manifold N is 13-dimensional. The condition that Z be tangent to N , $Z \lrcorner d\Sigma^a = 0$ and $Z \lrcorner dT^A = 0$, allows the components ξ^r , Ψ^A to be determined up to the factor μ and Φ^A is still given in 4.6(5).

The pair $(N, \iota^* \omega)$ is seen to be a contact manifold (in the sense of §4.4c) since $p \to Z(p)$ is the 1-dimensional characteristic distribution on N defined by $\iota^* \omega$. According to [26] the unparameterized integral curves of Z represented by $\lambda \to p(\lambda)$ coincide with the leaves of the foliation $p \to Z(p)$ of $\ker(\iota^* \omega)$ on N . The curves $\lambda \to \iota(p(\lambda))$ in $TG(V)$ are contained in the foliation defined by $\ker \omega$ on $TG(V)$ and the two quotient manifolds $TG(V)/\ker \omega$ and $N/\ker(\iota^* \omega)$ are diffeomorphic (we are assuming in the first place that these quotient sets are manifolds).

However, our interest is in the submanifold $C \subset TG(V)$ which is the intersection of the constraint manifolds M (5.4(2) and 5.4(3)), and N ((3), (4), (5)). The next section carries out the constraint procedure outlined in §5.3 for $C = M \cap N \subset TG(V)$.

§5.6 Calculation of the Constraint Equations.

The constraint formalism defined by 5.3(5) becomes, in the context of §5.5:

(1) $\Sigma^o = 0$

(2) $\Sigma^A = 0$

(3) $T^A = 0$

(4) $\phi^\Gamma = 0$

(5) $\bar{\phi}^\Gamma = 0$

(6) $Z \lrcorner d\Sigma^o = 0$

(7) $Z \lrcorner d\Sigma^A = 0$

(8) $Z \lrcorner dT^A = 0$

(9) $Z \lrcorner d\phi^\Gamma = 0$

(10) $Z \lrcorner d\bar{\phi}^\Gamma = 0$

and

(11) $Z \lrcorner \omega = \alpha d\Sigma^o + \beta_A d\Sigma^A + \mu_A dT^A + \rho_\Gamma d\phi^\Gamma + \sigma_\Gamma d\bar{\phi}^\Gamma$.

Equations (1) to (5) we call the <u>constraint equations</u>, (6) to (10) the <u>compatibility conditions</u>, and (11) the <u>foliation equations</u>. The task of this section is to evaluate the constraint equations, to calculate (6) to (10) subject to the constraints and to write out (11). A few special solutions are worked out in the following section while the general solution is presented in §5.8.

Explicitly the constraint equations are:

(12) $$v^o = 1$$

(13) $$Mv_A + \nu\, S_{AB}\, U^B = 0$$

(14) $$MU_A - v^r\, R_{rA} = 0$$

(15) $$\phi^\Gamma = 0$$

(16) $$v^A - g^A = 0 \quad .$$

Calculation of the differentials $d\Sigma^o$, $d\Sigma^A$, dT^A , $d\phi^\Gamma$, $d\bar{\phi}^\Gamma$ is straightforward although tedious. The results are:

(17) $$d\Sigma^o = dv^o$$

(18) $$d\Sigma_A = Mdv_A + \nu(v_A S_B + \epsilon_{AK}{}^L\, U^K\, I_{LB})d\Omega^B + \nu\, S_{AB}\, dU^B$$

(19) $$dT_A = -\frac{1}{2}\, S^{MN} R_{MNrA|b}\, v^r \theta^b + g_{AB}\tilde{E}^B + f_{AB}\, \tilde{F}^B$$
$$+ R_{Ab}\, dv^b + (\nu U_A S_B - \frac{1}{2}\, \epsilon^{MNC} I_{CB} R_{MNrA} v^r)d\Omega^B + MdU^A$$

(20) $$d\phi^\Gamma = F^\Gamma_a\, \theta^a$$

(21)
$$d\bar{\phi}^{\Gamma} = v^r \, F^{\Gamma}_{r|s}\theta^s - F^{\Gamma}_{K} v_L \, \epsilon^{KL}{}_{B} \, \tilde{E}^B$$
$$+ (F^{\Gamma}_{B} + \nu \, f^r v_B)\tilde{F}^B + F^{\Gamma}_{r} \, dv^r \quad .$$

In equation (19) the shorthand notation:

(22)
$$f_{AB} \equiv R_{AB} - \nu(S_B{}^N R_{oNrA} v^r + R_{oA} v_B)$$

(23)
$$g_{AB} \equiv S^M R_{MBrA} v^r - R_{AK} v^L \, \epsilon^K{}_{LB} + MU_K \, \epsilon^K{}_{AB}$$

is used.

If

$$Z = \nu^a \underset{\sim}{e}_a + P^A \underset{\sim}{E}_A + Q^A \underset{\sim}{F}_A + \xi^a \frac{\partial}{\partial v^a} + \Phi^A \frac{\partial}{\partial \Omega^A} + \Psi^A \frac{\partial}{\partial U^A}$$

denotes a generic vector field on $TG(V)$, then use of equations (17) to (21) gives for the compatibility equations (6) to (10):

(24)
$$\xi^o = 0$$

(25)
$$M\xi_A + \nu(v_A S_B + \epsilon_{AK}{}^L U^K I_{LB})\Phi^B + \nu \, S_{AB} \, \Psi^B = 0$$

(26)
$$- \frac{1}{2} \, S^{MN} R_{MNrA|b} v^r \nu^b + g_{AB} P^B + f_{AB} Q^B + R_{AB} \, \xi^B$$
$$+ (\nu U_A S_B - \frac{1}{2} \, \epsilon^{MNC} I_{CB} R_{MNrA} v^r)\Phi^B + M \, \Psi^A = 0$$

(27)
$$\nu^A - \nu^o \, g^A = 0$$

(28)
$$\xi^A + (\delta^A_B - \nu \, v^A v_B)Q^B + \epsilon^A{}_{BC} P^B v^C + \tilde{G}^A = 0$$

where

(29) $$\tilde{G}^A \equiv G^A_\Gamma \, F^\Gamma_{r|s} \, v^r \nu^s \quad .$$

Note that (16) and (27) together with (12) imply that

(30) $$\nu^a = \nu^o \, v^a$$

and

(31) $$f^\Gamma = -F^\Gamma_A \, v^A \quad .$$

We use the expression 4.6(2) for $Z \lrcorner \, \omega$ together with the differentials (17) through (21) to obtain from the foliation equations (11):

(32) $$S_A \Phi^A = -\frac{1}{2} S^{MN} R_{MNrA|o} \, v^r \mu^A - v_A P^A + \sigma_\Gamma \, F^\Gamma_{r|o} \, v^r$$

(33) $$M(\nu^o U_A - Q_A) = -\frac{1}{2} S^{MN} R_{MNrB|A} v^r \mu^B + \rho_A + \sigma_\Gamma \, F^\Gamma_{f|A} \, v^r$$

(34) $$I_{AB} \Phi^B + S_{AB} P^B = -\mu^B g_{BA} - \epsilon_{ABC} v^B \, \sigma^C$$

(35) $$\nu^o M \, v_A + \nu \, S_{AB} Q^B = \mu^B f_{BA} + (\delta^B_A - \nu \, v^B v_A) \sigma_B$$

(36) $$P^A - \nu^o \Omega^A = \nu(\Omega^A v_B + \epsilon^A{}_{BC} U^C) \beta^B$$
$$+ (\nu \Omega^A U_B - \frac{1}{2} \epsilon^{AMN} R_{MNrB} v^r) \mu^B$$

(37) $$\alpha - R_{oA} \, \mu^A - \sigma_A v^A = 0$$

(38) $$M\beta_A + \sigma_A - R_{AB} \, \mu^B = 0$$

(39) $$M\mu_A + \nu \, \beta^B \, S_{BA} = 0 \quad .$$

The notation $\sigma_A = F_A^\Gamma \sigma_\Gamma$, $\mu_A = F_A^\Gamma \mu_\Gamma$, $\beta_A = F_A^\Gamma \beta_\Gamma$ and $\rho_A = F_A^\Gamma \rho_\Gamma$ has been used in (32) through (39).

At this point, the necessary equations have their explicit form: we must solve (32) through (39) in conjunction with (24) to (28) and (12) to (16).

Before concluding this section we derive a number of relations for the Lorentz case (ν=1). Equation (39) implies that

$$\beta^A = -Ms^{-2} S^{AB}\mu_B + s^{-2}(\beta_B S^B)S^A$$

where

(40) $$s^2 = S_A S^A = \frac{1}{2} S_{AB} S^{AB} \quad .$$

Note also (for both the Galilei and Lorentz cases) that (13) gives

(41) $$S_A v^A = 0 \quad .$$

Multiplying S_A into (35) together with the use of the definition (22) and equations (38) and (41) result in

(42) $$S_A \beta^A = 0 \quad .$$

Therefore,

(43) $$\beta^A = -Ms^{-2} S^{AB} \mu_B \qquad (\text{if } \nu=1) \quad .$$

It is a straightforward computation that

(44) $$v^A \beta_A = U^A \mu_A \quad .$$

If we add (37) to $v_A \cdot$ (38) and use (44) we get

(45) $$\alpha = 0 \quad .$$

§5.7 Special Cases.

(a) V is a Newtonian spacetime; we take the equations derived in the last section and set $\nu = 0$ to obtain:

Constraint equations

(1) $$v^o = 1$$

(2) $$v^A = 0$$

(3) $$MU_A - R_{oA} = 0$$

(4) $$\phi^\Gamma = 0$$

(5) $$f^\Gamma = 0$$

Compatibility equations

(6) $$\xi^a = 0$$

(7) $$-\frac{1}{2}\nu^o S^{MN} R_{MNoA|o} + g_{AB}P^B + f_{AB}Q^B - \frac{1}{2}\Phi^B I_{BC}\epsilon^{CMN}R_{MNoA} + M\Psi_A = 0$$

(8) $$Q^A = -\nu^o G^A_\Gamma F^\Gamma_{o|o}$$

Foliation equations

(9) $\mu_A = \beta_A = \sigma_A = 0 = \alpha$

(10) $P^A = \nu^o\, \Omega^A$

(11) $\phi^A = -\nu^o\, \tilde{I}^{AB}\, S_{BC}\Omega^C$

(12) $\rho^A = \nu^o(R_o{}^A + G^A_\Gamma\, F^\Gamma_{o|o})$

The main feature to notice here is that since $\alpha = 0 = \beta_A = \mu_A$, there are no forces constraining the system to N. This essentially means that the condition that Z be a second-order equation is automatically satisfied in the Galilei case.

The forces constraining the motion to M are given by (12) which is the same as 5.4(16) if one remembers the definition 5.5(6).

Since $\nu^A = \nu^o v^A = 0$, we again have $v^\alpha = \dfrac{dx^\alpha}{d\lambda} = Z \lrcorner\, dx^\alpha = \nu^o e^\alpha_o$ (5.4(25)). That this is compatible with $\dot{e}^\alpha_o = e^\alpha_A\, Q^A$ (obtained from $\dfrac{de^\alpha_o}{d\lambda} = Z \lrcorner\, de^\alpha_o$) can be seen by explicitly calculating (8) in terms of the definition of ϕ^Γ, 5.4(1), and then comparing with the derivative of $v^\alpha = \nu^o e^\alpha_o$.

(b) V is Minkowski space; the conditions $\nu = 1$ and $R_{abcd} = 0 = R_{abcd|k}$ are used in §5.6 to obtain:

Constraint equations

(13) $v^o = 1$

(14) $v^a = 0$

(15) $U^A = 0$

(16) $\phi^\Gamma = 0$

(17) $f^\Gamma = 0$

Compatibility equations

(18) $\xi^a = 0$

(19) $\Psi^A = 0$

(20) $Q^A = -\nu^o \, G^A_\Gamma \, F^\Gamma_{o|o}$

Foliation equations

(21) $\sigma_A = -\nu^o \, S_{AB} \, G^B_\Gamma \, F^\Gamma_{o|o}$

(22) $\beta_A = -M^{-1} \, \sigma_A$

(23) $\mu_A = -M^{-1} \, \beta^B S_{BA} = -\nu^o M^{-2} \, S_{AB} S^B{}_C G^C_\Gamma \, F^\Gamma_{o|o}$

(24) $P^A = \nu^o \, \Omega^A$

(25) $\Phi^A = -\nu^o \, \tilde{I}^{AB} \, S_{BC} \Omega^C$.

α was shown to be zero at the end of §5.6, while for ρ_A we find

(26) $$\rho_A = \nu^o \, M(\delta_{AB} - S_{CB} \, G^C \, F^\Gamma_{o|A}) G^B_\Lambda \, F^\Lambda_{o|o} \quad .$$

Unlike for the Galilei situation, the constraint forces α, β^A , μ^A are non-zero in general in the Lorentz case. The formalism

of §5.4 does not take into account the conditions that the motions of the system satisfy a second-order principle; hence, the equations derived in that case did not fully determine the motion.

The integral curves follow as in the Galilei situation; the equations for $\dot{\Omega}^{\alpha}$ and $\dot{S}^{\alpha}$ are the same as in §4.6 while those for v^{α} and $\dot{e}^{\alpha}_{o}$ are given a priori by the constraints:

$$v^{\alpha} = \frac{dx^{\alpha}}{d\lambda} = Z \rfloor dx^{\alpha} = \nu^{o} e^{\alpha}_{o} \quad .$$

(c) V is DeSitter space time (or, a space of constant curvature).

If K denotes the constant curvature ([53]), then

(27) $$R_{\alpha\beta\gamma\delta} = K(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma})$$

and K is related to the Ricci scalar by

(28) $$K = -\frac{1}{12} R \quad .$$

The components of the curvature tensor in a frame (e_a) are

(29) $$R_{abcd} = K(\eta_{ac}\eta_{bd} - \eta_{ad}\eta_{bc})$$

where

(30) $$\eta = \mathrm{diag}(-1,1,1,1) \quad .$$

Note that

(31) $$R_{abcd|k} = 0 \quad .$$

Recall (§5.5(6)) that

(32) $$R_{k\ell} = \frac{1}{2} S^{MN} R_{MNk\ell} = K S^{MN} \delta_{Mk} \delta_{N\ell} \quad .$$

Using this in §5.6(14) gives

(33) $$U_A = -M^{-1} K S_{AB} v^B \quad .$$

Substitution of (33) into 5.6(13) yields the condition

(34) $$(M^2 + K s^2) v_A = 0 \quad .$$

which we rewrite as

(35) $$(1 + M^{-2} s^2 K) v_A = 0 \quad .$$

At this time recall the discussion following 4.6(12); equation 4.6(12) evaluated in the current context is precisely

(36) $$\Delta = 1 + M^{-2} s^2 K$$

which is assumed to be non-zero, so that

(37) $$v^A = 0 \quad .$$

Although not evident at this point, we have come back to the same situation as in Minkowski space (example (b)). This is related to the fact that although V is curved, the curvature is completely isotropic and we expect no net gravitational force to act

on the body. It may also be noted that $K = 0$ is just Minkowski space, example (b).

We take up the calculations again by examining the compatibility equations §5.6(25, 26, 28). In this case, they are:

(38) $$M\xi^A + S^A{}_B\Psi^B = 0$$

(39) $$KS^A{}_B\xi^B + M\Psi^A = 0$$

(40) $$\xi^A + Q^A + \tilde{G}^A = 0 \quad .$$

Substitution of (39) and (40) into (38), and noting that (38) implies $S_A\xi^A = 0$, we get

$$(M^2 + Ks^2)(Q^A + \tilde{G}^A) = 0 \quad ,$$

hence

(41) $$Q^A = -\tilde{G}^A = -\nu^o \, G^A_\Gamma \, F^\Gamma{}_{o|o} \quad .$$

Next, because $v^A = 0$, we get:

(42) $$\nu^A = 0$$

(43) $$U^A = 0 \quad ,$$

and it turns out that

(44) $$f_{AB} = 0$$

and

(45) $$g_{AB} = 0 \ .$$

With these results, it is easy to show that again

(46) $$P^A = \nu^o \ \Omega^A$$

and

(47) $$\Phi^A = -\nu^o \ \tilde{I}^{AB} \ S_{BC} \ \Omega^C \ .$$

(d) Summary. We have examined in three examples of space-times the motion of a rigid body whose centre of mass is constrained to a given worldline.

In all of these cases, constraining the motion of the centre of mass to the worldline has no effect on the rotational motion of the body. This is because equations (11), (25) and (47) which determine Φ^n for each example are exactly the same as 4.6(5) and results in the unconstrained equations of motion 4.6(30) for Ω and 4.6(31) for S:

$$\bar{\delta}^\alpha_\beta \ \dot{\Omega}^\beta = -\nu^o \ \tilde{I}^{\alpha\beta} \ S_{\beta\gamma} \ \Omega^\gamma$$
$$\bar{\delta}^\alpha_\beta \ \dot{S}^\beta = 0.$$

Again, both M and S are conserved. (see Remark (iii) in §4.6c.)

The fundamental difference between the Newtonian example and the relativistic examples is the appearance in the latter of constraint forces associated to the submanifold N (§5.5b), but not in the former.

§5.8 General Solution of the Relativistic Constraint Problem.

We return to the general setting of §5.6 and set $\nu = 1$ to examine the general Lorentz case. A glance at the foliation equations and in particular 5.6(38) and 5.6(39) shows that most components of the vector field Z and the constraints α, β_A, etc. can be expressed in terms of μ_A. Accordingly, in this section we algebraically derive a system of equations which determine μ_A in terms of the coordinates on $T\,Lor(V)$ and the constraints ϕ^Γ. The derivation is horrendously complicated in detail so we present only the highlights of the calculations.

In addition to equations 5.6(12) to 5.6(16) and 5.6(22) through to 5.6(45), we need the facts:

(1) $$U_A v^A = 0$$

(2) $$S_A \mu^A = 0 \quad .$$

The latter equation shows that the problem is essentially two dimensional - we can express μ_A as, for example, $\mu_A = \alpha v_A + \beta S_{AB} v^B$. The following abbreviations are used throughout:

(3) $$\lambda = S_A \Omega^A$$

(4) $$\gamma = S_A U^A$$

(5) $$\eta = U_A \Omega^A \quad .$$

The substitution of 5.6(43) into 5.6(38) gives

(6) $$\sigma_A = (R_{AB} + M^2 s^{-2} S_{AB})\mu^B \quad .$$

Equation 5.6(36) yields after some calculation

(7) $$P_A = \nu_o \Omega_A + h_{AB}\mu^B$$

where

(8) $$h_{AB} = 2\,\Omega_A U_B + M\,s^{-2}\gamma\delta_{AB} - M\,s^{-2}\,S_A U_B - \frac{1}{2}\,\epsilon_A{}^{MN} R_{MNrB} v^r \quad .$$

Substitution of (6), (7) and 5.6(23) into 5.6(34) eventually yields

(9) $$I_{AB}\Phi^B = -\nu^o\,S_{AB}\Omega^B + k_{AB}\mu^B$$

with

(10) $$k_{AB} = -S_A{}^C h_{CB} - S^M R_{MArB} v^r + MU^K\,\epsilon_{KAB} + M^2 s^{-2}\,S_A v_B \quad .$$

If we now multiply 5.6(35) by S_{AB} (taking note of the fact that

$S_{AC}S^C{}_B = S_A S_B - s^2\,\delta_{AB}$) and make the appropriate subsitutions: 5.6(22), and (6) being the major ones we get (eventually)

(11) $$Q_A = \nu^o M s^{-2}\,S_{AB}v^B - s^{-2}(S_K\tilde{G}^K)S_A + \ell_{AB}\,\mu^B$$

with

(12) $$\ell_{AB} = s^{-2}\,S_A v^s S^N R_{sNrB}\,v^r - R_{oArB}\,v^r + M^2 s^{-2}\,\delta_{AB} - s^{-2}\gamma\,S_A U_B \quad .$$

Now we focus our attention on the compatibility equations 5.6(25) to 5.6(28). The idea is essentially the same as that used to derive 5.7(41): equations 5.6(28) and 5.6(26) are substituted into 5.6(25) and from this we derive the desired relation for μ_A .

First, though, we simplify 5.6(28) by making appropriate substitutions and reductions. It turns out that

(13) $$\xi_A = -\nu^o M s^{-2} S_{AB} v^B + s^{-2}(S_K \tilde{G}^K) S_A + \nu^o \epsilon_{ABC} v^B C - \tilde{G}_A + p_{AK} \mu^K$$

where

(14) $$p_{AK} = -(\delta_A^B - v_A v^B)\ell_{BK} + \epsilon_{AB}{}^C v^B h_{CK} \quad .$$

At this point, all the intermediate quantities of interest, for example P^A , Q^A , have been expressed in terms of μ_A . After an extremely laborious computation, the equation for μ_A can be written in the form:

(15) $$T_{AK} \mu^K = L_A \quad .$$

(T_{AK} and L_A are written out below.) Recall that $\mu_A S^A = 0$; this implies that there exists α and β such that

(16) $$\mu_A = \alpha v_A + \beta S_{AB} v^B \quad .$$

After we give T_{AK} and L_A , we will calculate α and β .

$$(17)\qquad T_{AK} = (M^2U^2 - M^4s^{-2})\delta_{AK} + M^2s^{-2}\gamma S_A U_K$$

$$- 2M^2s^{-2}\gamma v_A \Omega^C S_{CK} + 2M^2 v_A U^B \Omega^C \epsilon_{BCK}$$

$$+ (2M^3s^{-2}\lambda - 2M^4s^{-2})v_A v_K - \frac{1}{2} M\gamma \epsilon_A{}^{MN} R_{MNrK} v^r$$

$$+ S_A{}^B \{-M^2s^{-2} S_B{}^C v^s R_{sCrK} v^r$$

$$- R_{oNrB} v^r S^{NC}(M^2s^{-2}\delta_{CK} - R_{oCsK} v^s)$$

$$+ \frac{1}{2} S^L R_{LCsB} v^s \epsilon^{CMN} R_{MNrK} v^r - M U^M R_{MKrB} v^r\}$$

and

$$(18)\qquad L_A = \nu^o M^3s^{-2}S_{AB}v^B - M^2 s^{-2}(S_K\tilde{G}^K)S_A$$

$$+ M^2\tilde{G}^A + S_{AB}\tilde{R}^B + \nu^o M v^N R_{oNrB} v^r S^B{}_A$$

$$- S_{AB} R^{BC} \tilde{G}_C \quad ,$$

where we have put

$$(19)\qquad \begin{cases} \tilde{R}_A \equiv -\frac{1}{2} S^{MN} R_{MNrA|b} v^r \nu^b \\ U^2 \equiv U_A U^A \\ v^2 \equiv v_A v^A \quad . \end{cases}$$

If (16) holds, it must satisfy (15):

$$(20)\qquad \alpha T_{AK}v^K + \beta T_{AK} S^{KU} v_U = L_A \quad .$$

By contracting (20) with v^A and $S^{AT} v_T$ we obtain the system of equations:

$$
(21) \qquad \begin{cases} A\,\alpha + B\,\beta = \overline{M} \\ C\,\alpha + D\,\beta = \overline{N} \end{cases}
$$

where we have put:

$$
(22) \qquad \begin{cases} A = v^{A}\; T_{AK}\; v^{K} \\ B = v^{A}\; T_{AK}\; S^{KU}\; v_{U} \\ C = S^{AT}\; v_{T}\; T_{AK}\; v^{K} \\ D = S^{AT}\; v_{T}\; T_{AK}\; S^{KU}\; v_{U} \\ \overline{M} = v_{A}\; L^{A} \\ \overline{N} = L^{A}\; S_{AT}\; v^{T} \qquad . \end{cases}
$$

Then:

$$
(23) \qquad \alpha = \overline{\Delta}^{-1}(MD-NB)
$$

$$
(24) \qquad \beta = \overline{\Delta}^{-1}(AN-CM)
$$

with

$$
(25) \qquad \overline{\Delta} = AD - BC \qquad .
$$

The expressions for the quantities (22) follow next.

(26)
$$A = M^2 s^{-2}[\gamma^2 - M^2(1+v^2)]v^2 + \gamma\, U^M S^N R_{MNoB}\, v^B$$
$$+ M^2 v^B\, R_{oBoC}\, v^C$$
$$+ v_A S^{AB}\, \{-R_{oNrB} v^r\, S^{NC}(M^2 s^{-2} v_C - R_{oCoK}\, v^K)$$
$$+ \frac{1}{2}\, S^L R_{LCsB}\, v^s\, \epsilon^{CMN} R_{MNoK}\, v^K$$
$$- M\, U^M\, v^N\, R_{MNrB}\, v^r\}$$

(27)
$$B = -\frac{1}{2} M \gamma v_A\, \epsilon^{AMN} R_{MNrK}\, v^r\, S^{KL}\, v_L$$
$$+ v_A s^{AB}\, \{-M^2 s^{-2} S_B{}^C\, v^s\, R_{sCrK}\, v^r$$
$$- R_{oNrB}\, v^r S^{NC}(M^2 s^{-2} \delta_{CK} - R_{oCsK}\, v^s)$$
$$+ \frac{1}{2}\, S^L R_{LCsB}\, v^s\, \epsilon^{CMN} R_{MNrK}\, v^r$$
$$- M\, U^M\, R_{MKrB}\, v^r\}\, S^{KU}\, v_U$$

(28)
$$C = M(\gamma\, S^M - s^2\, U^M) v^N\, R_{MNoB}\, v^B$$
$$- M^2\, v_B S^{BC}\, v^s\, R_{sCoK}\, v^K$$
$$- R_{oNoB}\, v^B\, S^{NC}(M^2 v_C - s^2 R_{oCoK}\, v^K)$$
$$+ \frac{1}{2}\, s^2 S^L\, R_{LCoB}\, v^B\, \epsilon^{CMN}\, R_{MNoK}\, v^K$$

(29) $$D = M^2 v^2 (s^2 U^2 - M^2)$$

$$+ \{M\gamma S^M v^N R_{MNrK} v^r - M^2 v_B S^{BC} v^s R_{sCrK} v^r$$

$$- R_{oNoB} v^B S^{NC} (M^2 \delta_{CK} - s^2 R_{oCsK} v^s)$$

$$+ \frac{1}{2} s^2 S^L R_{LCoB} v^B \epsilon^{CMN} R_{MNrK} v^r$$

$$- s^2 M U^M R_{MKoB} v^B\} S^{KU} v_U$$

(30) $$\overline{M} = M^2 v_A \tilde{G}^A + v_A S^{AB} (\tilde{R}_B - \nu^o M v^N R_{oNrB} v^r - R_{BC} \tilde{G}^C)$$

(31) $$\overline{N} = -M(S_K \tilde{G}^K)\gamma + M s^2 U_A \tilde{G}^A$$

$$+ s^2 v_B (\nu^o M^3 s^{-2} v^B + \tilde{R}^B - \nu^o M v^N R_{oNrB} v^r - R^{BC} \tilde{G}_C) \quad .$$

Finally, note that

(32) $$S_A L^A = 0 \quad .$$

It is not practical to write out (23), (24) and (25) in the general situations. What we have in mind is that for a particular problem (22) is first calculated, then (25) and finally (23) and (24); (16) then gives μ_A and the other quantities such as Q^A (11) will follow.

We therefore have an explicit algorithm for calculating the constrained motion of a rigid test body in a general curved space-time. An example of this is given in the next section; however, we will only be calculating an approximate solution.

§5.9 Motion Along a Schwarzschild Circular Orbit.

(a) The examples presented in §5.7 all lead to the result that constraining to a worldline (for those special cases) does not introduce any new features of the motion, namely, a precession of the spin axis. This was expected for a Newtonian or Minkowskian space-time and could be reasonably justified for a DeSitter universe. We turn now to consider an example in a Schwarzschild space-time with the hope that there is enough anisotropy present to produce some sort of effect (a Schwarzschild space-time admits a four-parameter group of isometries compared to a ten-dimensional group for DeSitter and Minkowski spaces).

The example is that of a rigid (test) body constrained at its centre of mass to move along a circular orbit about the source producing the Schwarzschild field; one may think, for example, of a gyroscope orbiting the earth. We first give some details about such orbits then find an approximate solution of the equations of §5.8 and finally calculate the spin precession.

(b) Schwarzschild orbits. We shall employ Schwarzschild coordinates ([35]; or curvature coordinates [53]) throughout. In this coordinate system, (t,r,θ,ϕ), the Lorentz metric takes the form

$$(1) \qquad g \equiv \mathrm{diag}\left(-\left[1 - \frac{2r_o}{r}\right], \left[1 - \frac{2r_o}{r}\right]^{-1}, r^2, r^2 \sin^2\theta\right)$$

where $2r_o$ is the gravitational radius corresponding to the mass of the object producing the field.

It is clear from (1) that $\frac{\partial}{\partial t}$ and $\frac{\partial}{\partial \phi}$ are Killing vectors of g (generating time translations and ϕ - rotations respectively). Denote by E ("energy at infinity") and L ("angular momentum") the corresponding conserved quantities, and let

$$(2) \qquad \begin{cases} \hat{E} \equiv \dfrac{E}{M} \\ \\ \hat{L} \equiv \dfrac{L}{M} \quad . \end{cases}$$

Orbital motion takes place in a plane; choose this plane to be

$$(3) \qquad \theta = \frac{\pi}{2} \quad .$$

The equations of the orbit are then

$$(4) \qquad \left(\frac{dr}{d\tau}\right)^2 = \hat{E}^2 - \tilde{V}^2(r)$$

with

$$(5) \qquad \tilde{V}(r) = \left[\left(1 - \frac{2r_o}{r}\right)\left(1 + \frac{\hat{L}^2}{r^2}\right)\right]^{1/2} \quad ,$$

$$(6) \qquad \frac{d\phi}{d\tau} = \frac{\hat{L}}{r^2}$$

and

$$(7) \qquad \frac{dt}{d\tau} = \frac{\hat{E}}{1 - \dfrac{2r_o}{r}} \quad .$$

$\tilde{V}(r)$ defined by (5) is an "effective potential" for the orbital motion of the test body. The curve parameter τ is proper time for the body.

Stable circular orbits occur at the minimum of $\tilde{V}$ and are only possible for $\hat{L} > \frac{2}{\sqrt{3}} r_o$; for example

(8) $$R = \frac{6r_o}{1 - (1 - 12r_o/\hat{L}^2)^{1/2}}$$

gives the value of r for a body moving with angular momentum L in a circular orbit. Taking $r = R$ in (6) and (7) determines ϕ and t as linear functions of τ .

We shall therefore confine our attention to the orbit

(9) $$\tau \to (k\tau, R, \tfrac{\pi}{2}, \ell\tau)$$

where

(10) $$\begin{cases} k = \dfrac{\hat{E}}{1 - \dfrac{2r_o}{R}} \\[2ex] \ell = \dfrac{\hat{L}}{R^2} \quad , \end{cases}$$

and if one defines:

(11) $$q = \frac{\ell}{k}$$

(12) $$\begin{cases} \phi^1 = x^1 - R \\ \phi^2 = x^2 - \frac{\pi}{2} \\ \phi^3 = x^3 - qx^o \quad , \end{cases}$$

the orbit is also described by $\{x \in V : \phi^\Gamma(x) = 0\}$.

From (12) we find (cf. 5.4(7))

$$(13)\quad \begin{cases} F^{\Gamma}_{\alpha} = \dfrac{\partial\phi^{\Gamma}}{\partial x^{\alpha}} = \delta^{\Gamma}_{\alpha} - q\,\delta^{\Gamma}_{3}\,\delta^{o}_{\alpha} \\ F^{\Gamma}_{a} = e^{\Gamma}_{a} - q\,\delta^{\Gamma}_{3}\,e^{o}_{a} \\ F^{\Gamma}_{\alpha|\beta} = -\Gamma^{\Gamma}_{\alpha\beta} + q\,\delta^{\Gamma}_{3}\,\Gamma^{o}_{\alpha\beta} \\ F^{\Gamma}_{a|b} = e^{\alpha}_{a}\,e^{\beta}_{b}\,F^{\Gamma}_{\alpha|\beta} \quad . \end{cases}$$

The tangent field to (9) is

$$(14)\quad \tau \to (v^{\alpha}(\tau)) = (k,0,0,\ell) \quad .$$

The nonvanishing Christoffel symbols are

$$(15)\quad \begin{cases} \Gamma^{0}_{10} = r_o r^{-2} e^{\alpha} \quad , & \Gamma^{1}_{00} = r_o r^{-2} e^{-\alpha} \\ \Gamma^{1}_{11} = -r_o r^{-2} e^{\alpha} \quad , & \Gamma^{1}_{22} = -re^{-\alpha} \\ \Gamma^{1}_{33} = -re^{-\alpha}\sin^2\theta \quad , & \Gamma^{2}_{12} = r^{-1} \\ \Gamma^{2}_{33} = -\sin\theta\cos\theta \quad , & \Gamma^{3}_{13} = r^{-1} \\ \Gamma^{3}_{23} = \cot\theta & \end{cases}$$

where

$$(16)\quad e^{-\alpha} = 1 - \frac{2r_o}{r}$$

has been introduced for brevity. Next, the nonvanishing coordinate components of the Riemann curvature tensor are:

$$
(17)\qquad \begin{cases}
R^1{}_{010} = -2r_o r^{-3} e^{-\alpha} \\
R^1{}_{212} = 2r_o r^{-1} \\
R^2{}_{020} = r_o r^{-3} e^{-\alpha} \\
R^2{}_{323} = 2r_o r^{-1} \sin^2\theta \\
R^3{}_{030} = r_o r^{-3} e^{-\alpha} \\
R^3{}_{131} = 2r_o r^{-3} e^{\alpha} \quad .
\end{cases}
$$

However, of more importance are the curvature components $R^a{}_{bcd}$ in a Lorentz frame (e_a). First of all, in the natural static orthonormal frame

$$
(e^{\alpha}_{*a}) = \mathrm{diag}\left(e^{\frac{\alpha}{2}},\ e^{-\frac{\alpha}{2}},\ r^{-1},\ r^{-1}(\sin\theta)^{-1}\right),
$$

the curvature components are:

(21)
$$- \underset{*}{R}{}^{1}{}_{010} = 2r_0 r^{-3} = \underset{*}{R}{}^{1}{}_{212} = \underset{*}{R}{}^{2}{}_{323} = \underset{*}{R}{}^{3}{}_{131}$$
$$\underset{*}{R}{}^{2}{}_{020} = r_0 r^{-3} = \underset{*}{R}{}^{3}{}_{030}$$

A general frame (e_a) is obtained from $(\underset{*}{e}_a)$ by a Lorentz transformation $(L^a{}_b)$:

(22)
$$e_a = \underset{*}{e}_b \, L^b{}_a \quad \text{and} \quad \theta^a = \tilde{L}^a{}_b \, \underset{*}{\theta}{}^b$$

where

$$\tilde{L} \equiv (L^{-1})^t .$$

Then for components in the general frame (e_a),

(23)
$$R^a{}_{bcd} = \tilde{L}^a{}_k \, \underset{*}{R}{}^k{}_{\ell mn} \, L^\ell{}_b \, L^m{}_c \, L^n{}_d \, .$$

(b) Basis for approximation. In terms of any suitable norm $||\cdot||$ we see from (21) that

(24)
$$||\underset{*}{R}{}^a{}_{bcd}|| \sim r_0 r^{-3} .$$

The extension of L to $\tilde{L} \otimes L \otimes L \otimes L$ acting on the right on the space of (1,3)-tensors has norm

(25)
$$||\tilde{L} \otimes L \otimes L \otimes L|| = ||\tilde{L}^a{}_k \, L^\ell{}_b \, L^m{}_c \, L^n{}_d|| = ||L||^2$$

so that from (23) and (24) we obtain

(26)
$$||R^a{}_{bcd}|| \sim ||L||^2 \; r_0 r^{-3} \, .$$

The assumption is now made that for $r >> r_0$, $||L||$ is of the order 1

(meaning that the body is not moving too fast with respect to the source $r = 0$). Note that all considerations are made and, strictly speaking, valid here only in Schwarzschild coordinates.

To obtain a criterion for approximation, we first re-examine 4.6(21):

$$(27)\qquad \begin{cases} v^\alpha = \nu^o(e_o + \frac{1}{2} M^{-2}\Delta^{-1}S^{\mu\nu} R_{\mu\nu\gamma\delta} e_o^\gamma S^{\delta\alpha}) \\ \\ \Delta = 1 + \frac{1}{4} M^{-2} S^{\alpha\beta} R_{\alpha\beta\gamma\delta} S^{\gamma\delta} \quad . \end{cases}$$

It is now generally conceded ([12], [13], [27], [29]) that in a curved space-time momentum, $p^\alpha \equiv Me_o^\alpha$, need not be parallel to the velocity v^α, although this is more of a formal distinction than practical ([13]). In practical situations, the approximation is made whereby terms containing quadratic spin factors are neglected. However, we base our approximation on the assumption that

$$(28)\qquad \lambda \equiv \frac{\Delta - 1}{c^2} << 1.$$

Using (27) we have, independent of the frame e_a,

$$(29)\qquad \lambda = \frac{2M^{-2}s^2r_0}{r^3} \qquad (c = 1).$$

In the usual weak-field approximation, $r >> r_o$ and therefore (28) is satisfied provided the spin is not too large. (Note that the various discussions ([14], [4], [43]) of the satellite experiment assume λ is negligible.)

For the earth-gyroscope system mentioned earlier, assume that the gyroscope is a spherical 100 gm. mass with a 2 cm. radius and an angular velocity of $\omega = 10^3 \frac{\text{rad}}{\text{sec}}$, and that it is orbiting 500 kilometers above the earth. We temporarily revert to units in which the speed of light has its usual value: $c = 3\times10^{10}$ cm sec^{-1} . Then

$$s \sim 1.6\times10^5 \text{ gm cm}^2 \text{ sec}^{-1}$$

$$T_{rot} = \frac{s\omega}{c^2} \sim 1.7\times10^{-13} \text{ gm}$$

$$M \sim m = 100 \text{ gm} \quad R \sim 10^8 \text{ cm}$$

$$(r_o)_{earth} = .44 \text{ cm} ,$$

so that

$$\lambda = \frac{M^{-2} s^2 r_0}{c^2 R^3} \sim 10^{-37} .$$

Thus, for a reasonable physical situation λ is extremely small compared to unity.

(c) Approximate solution for μ_A . The first thing to note is that if the procedure used to solve §4.6(10) is applied to the system §5.6(13) and §5.6(14) we obtain

$$v^A = \frac{1}{2} M^{-2} \Delta^{-1} S^{MN} R_{MNoB} S^{BA} \tag{30}$$

where Δ is given in (27). This shows that both U^A and v^A are of the order λ .

Equations 5.8(26) to 5.8(3) to lowest order are:

$$(31)\quad \begin{cases} A = -M^4\, s^{-2}\, v^2 + O(\lambda^3) \\ B = 0 + O(\lambda^3) \\ C = 0 + O(\lambda^3) \\ D = -M^4\, v^2 + O(\lambda^3) \end{cases}$$

and

$$(32)\quad \begin{cases} \bar{M} = M^2 v_K \tilde{G}^K + v_K S^{KL}\, \tilde{R}_L + O(\lambda^2) \\ \bar{N} = M^2 (v_K S^{KL} \tilde{G}_L) + s^2 (v_K \tilde{R}^K) + O(\lambda^2) \quad . \end{cases}$$

Also,

$$(33)\quad \bar{\Delta} = AD = M^8 s^{-2} v^4 + O(\lambda^6) \quad .$$

Using these approximate values in 5.8(23) and 5.8(24) we get

$$(34)\quad \begin{cases} \alpha = -M^{-4} s^2 v^{-2} (M^2 v_A \tilde{G}^A + v_A S^{AB} \tilde{R}_B) \\ \beta = -M^{-4} v^{-2} (M^2 v_A S^{AB} \tilde{G}_B + s^2 v_A \tilde{R}^A) \quad . \end{cases}$$

Finally, substitution of (34) into 5.8(16) yields after some manipulation

$$(35)\quad \mu_A = M^{-2} S_{AB} (-S^{BC} \tilde{G}_C + M^{-2} s^2 \tilde{R}^B) + O(\lambda^2) \quad .$$

(Compare this with 5.7(23).)

We shall use (35) to give an estimate of the spin precession; first,

$$\frac{dS^\alpha}{d\lambda} = \frac{d}{d\lambda}(e^\alpha_A S^A) = S^A Z \lrcorner\, de^\alpha_A + e^\alpha_A\, I^A{}_B\, Z \lrcorner\, d\Omega^B \quad ,$$

whence

$$\dot{S}^\alpha = e^\alpha_o (S_A Q^A) - S_A{}^B P^A e^\alpha_B + e^\alpha_A\, I^A{}_B\, \Phi^B \quad ,$$

so that using the values 5.8(7, 9, 11),

$$(36) \qquad (\delta^\alpha_\beta + e^\alpha_o e_{\beta o})\dot{S}^\beta = e^\alpha_A (S^{AB} h_{BC} + k^A_C)\mu^C \quad .$$

Substitution of 5.8(8,10) together with (35) into (36) gives

$$(37) \qquad (\delta^\alpha_\beta + e^\alpha_o e_{\beta o})\dot{S}^\beta = (S^\mu R^\alpha_{\mu o\beta} + M\, \epsilon^\alpha{}_{\beta\gamma\delta}\, e^\Gamma_o U^\delta + M^2 S^{-2} S^\alpha\, v_\beta)$$

$$\times\, S^{\beta\rho}(M^{-2} S_{\rho\sigma}\tilde{G}^\sigma + M^{-4} S^2 \tilde{R}_\rho)$$

where

$$(38) \qquad \left\{ \begin{aligned} U_\alpha &= \frac{1}{2} M^{-1}\, S^{\mu\nu}\, R_{\mu\nu\rho\alpha}\, v^\rho \\ \tilde{G}^\alpha &= e^\alpha_A\, G^A_\Gamma\, F^\Gamma_{\rho|\sigma} v^\rho v^\sigma \\ \tilde{R}_\alpha &= -\frac{1}{2} S^{\mu\nu}\, R_{\mu\nu\rho\alpha|\sigma}\, v^\rho v^\sigma \quad . \end{aligned} \right.$$

In view of (38), (17) and (15),

$$(39) \qquad (\delta^\alpha_\beta + e^\alpha_o e_{\beta o})\dot{S}^\beta = 0 + \mathcal{O}(\lambda^2)$$

so that the spin precession is second order in λ , and therefore undetectable in any practical sense.

We now estimate the extent to which S^α undergoes Fermi-Walker transport along the orbit (recall §4.6e). The definition 4.6(45) of λ^2 is compatible with the definition (28) for λ $(c = 1)$ in orders of magnitude. Comparing (39) with 4.6(31) (they differ by $O(\lambda^2)$), equation 4.6(52) shows that Fermi-Walker transport of spin is first order in λ for the free case as well for constrained motion. Equation (31) says that for constrained motion, Fermi-Walker transport is given by 4.6(52) corrected by second-order terms.

Footnotes

1) Often we interpret a vectorfield $\xi \in \mathfrak{X}(M)$ as a linear differential operator of $C^\infty(M)$, the ring of all C^∞-functions on M.

2) In this discussion "e" no longer denotes the identity of the group.

3) A symplectic form is a closed 2-form which is non-degenerate at each point; a presymplectic form is a 1-form whose exterior derivative is a symplectic form.

4) Equation (33) shows that for free fall motion in curved space Euler's equations remain the same as in flat space. On the other hand, geodesic motion of the centre of mass is destroyed by the curvature of Σ. We meet an equation like (34) in §4.6.

5) This is a notational device only. In particular, a term including ν as a factor need not make sense in the Galilei case - it is merely ignored.

6) The more suggestive term "integrable" is beginning to replace "first order flat".

7) One often sees the definition that a contact structure is a 1-form θ such that $\theta \wedge (d\theta)^r$ is a volume on M where $r = \frac{1}{2}(\dim M - 1)$. In our context, θ is called an <u>exact contact structure</u> ([1]).

8) But more general is Souriau's model of a spinning particle in Minkowski space ([48], page 181). The evolution space is diffeomorphic to $R^4 \times R^2 \times S^2$ which is not T^*Q for any Q, and ω is not exact.

9) Chosen to be future-pointing with respect to the orientation.

10) Other than finite-dimensional.

11) Not to be confused with the term as used in the theory of ordinary differential equations.

12) A worldline is a 1-dimensional timelike submanifold of V.

GLOSSARY OF NOTATIONS

The following is a list of commonly used symbols and their meanings as used in the thesis. Reference to the text is given where appropriate.

General.

$\equiv$	definition
iff	if and only if; is equivalent to
$\subset$	$A \subset B$, A is contained in B
ε	$a \varepsilon A$, a is an element of A

Maps.

$f : A \to B$, $a \to f(a)$, f maps each $a \varepsilon A$ to $f(a) \varepsilon B$

$f(A)$	image of A under f
$f^{-1}(S)$	inverse image of $S \subset B : f^{-1}(S) = \{a \varepsilon A : f(a) \varepsilon S\}$
$f \circ g$	composition, g followed by f
f^{-1}	inverse map to f , $f \circ f^{-1} = id$

Linear Algebra.

V,W,...	finite-dimensional vector spaces
R	real numbers

$R^n = R \times \ldots \times R$, n times; prototype n-dimensional vector space

$(A^a{}_b)$	matrix whose (a,b)-entry is $A^a{}_b$
	Note: A^a_b , $A^a{}_b$, A_{ab} , $A_a{}^b$, A^{ab} , all denote an (a,b) entry

$\{e_a\}$	frame of V

components; $x = x^a e_a \in V$

V^*	dual space to V
$\{\theta^a\}$	frame of V^* dual to e_a
δ^a_b	Kronecker delta
A^t	transpose of A
$\langle\ ,\ \rangle$	natural pairing between V and V^*
$(\ ,\ )$	an inner product on V
$\otimes$	tensor product
$\wedge$	exterior product
$\lrcorner$	$X \lrcorner\, \omega$, interior product of the p-form ω by the vector X, yielding a (p-1) - form

<u>Lie Theory</u>.

G,H,...	Lie groups (§2.3e)
g,h,...	their Lie algebras
[,]	bracket operation in g
GL(n)	general linear group of $n\times n$ invertible real matrices (§2.3g)
$g\ell(n)$	Lie algebra of GL(n)
$\{E^a_b\}$	standard basis of $g\ell(n)$
SO(3)	orthogonal 3×3 matrices with positive determinant (§2.4a)
so(3)	Lie algebra of SO(3) (§2.4c)
Gal	homogeneous Galilei group (§4.2a)
Lor	connected component of identity of Lorentz group (§4.2b)

Manifolds.

C^∞	continuously differentiable to any order; smooth
$M, N, \ldots$	manifolds: C^∞, Hausdorff, paracompact, connected (§2.3a)
nbd	neighbourhood
(U, h)	local chart at $m \in M$ determining local coordinates (x^α)
$T_m M, T_m^* M$	tangent space, cotangent space at $m \in M$ (§2.3b)
$\pi : TM \to M$, $\pi : T^*M \to M$	tangent and cotangent bundles of M
$f, g \in C^\infty(M)$	real-valued smooth functions on M
$X \in \mathfrak{X}(M)$	smooth vector fields on M (§2.3a)
$\sigma \in \mathfrak{X}^*(M)$	smooth 1-forms on M
$\omega \in \mathfrak{X}^p(M)$	smooth p-forms on M
$\{\partial_\alpha \equiv \frac{\partial}{\partial x^\alpha}\}$	local $C^\infty(M)$ - module basis of $\mathfrak{X}(M)$
$\{dx^\alpha\}$	local $C^\infty(M)$ - module basis of $\mathfrak{X}^*(M)$

Maps and Operations on Manifolds.

$f : M \to N$	smooth mapping between manifolds
$f_* : TM \to TN$	induced differential map (§2.3b)
$f^* : T^*N \to T^*M$	induced dual map (§2.3b)
$f^* : \mathfrak{X}^p(N) \to \mathfrak{X}^p(M)$	pullback of p-forms
$d : \mathfrak{X}^p(M) \to \mathfrak{X}^{p+1}(M)$	exterior derivative of forms
L_X	Lie derivative by the vector field X
∇_α , $\vert\alpha$	covariant derivative in the ∂_α - direction
$\cdot \equiv v^\alpha \nabla_\alpha$	covariant derivative along a curve

Frame Bundles (also see the Appendix).

$L(M)$	bundle of linear frames of M
$SO(M)$	bundle of orthogonal frames (§3.2b)
$Lor(V), Gal(V)$	bundles of $\{Lor \mid Gal\}$ frames on space-time V (§4.3b)
$\omega^a{}_b$	g - valued connection 1-form on $G(V)$
$\underset{\sim}{A}$	fundamental (Killing) vector field associated to $A \in g$
(x^α, e^α_a)	coordinates on $L(M)$, $Lor(V)$, or $Gal(V)$
$\{\underset{\sim}{e}_a, \underset{\sim}{E}_A, \underset{\sim}{F}_A\}$	global basis of either $\mathfrak{X}(Gal(V))$ or $\mathfrak{X}(Lor(V))$

Miscellaneous.

L	a Lagrangian (§2.5b;§3.3a)
∇	3-dimensional gradient, $\nabla f \equiv (\partial_\Gamma f)$
m	(rest) mass of a body
T_{rot}, T_{trans}	rotational and translational kinetic energy (§3.3a)
M	mass of a spinning body, $M = m + \nu\, T_{rot}$ (§4.5d)
η	Minkowski metric, $\eta = \text{diag}(-1,1,1,1)$
g	Lorentz metric on space-time V
ϵ_{ABC}	3-dimensional alternating tensor
$\epsilon_{\alpha\beta\gamma\delta}$	Levi-Civita density
( )	symmetrization of indices
[ ]	antisymmetrization of indices

BIBLIOGRAPHY

[1] Abraham, R. and Marsden, J.E., Foundations of Mechanics, Benjamin, New York, 1967.

[2] Arnold, V., Sur la géometrie differential des groupes de Lie de dimension infinie et ses applications à l'hyrodynamique des fluides parfaits, Ann. Inst. Fourier, Grenoble 16, 319-361, (1966).

[3] Bailey, I. and Israel, W., Lagrangian dynamics of spinning particles and polarized media in general relativity, Comm. Math. Phys. 42, 65-82 (1975).

[4] Barker, B.M. and O'Connell, R.F., Nongeodesic motion in general relativity, Gen. Rel. Grav. 5, 539-544 (1974).

[5] Beiglböck, W.D., The centre-of-mass in Einstein's theory of relativity, Comm. Math. Phys. 5, 106-130 (1967).

[6] Cannon, R.H., Requirements and design for a special gyro for measuring general relativity effects from an astronomical satellite, article in IUTAM - Symposion Celerina 1962, Ziegler (ed.), Springer-Verlag, 1963.

[7] Chernoff, P.R. and Marsden, J.E., Properties of Infinite Dimensional Hamiltonian Systems, (Lect. Notes in Math. No. 425), Springer-Verlag, 1974.

[8] Corben, H.C., Classical and Quantum Theories of Spinning Particles, Holden-Day, San Francisco, 1968.

[9] Corinaldesi, E. and Papapetrou, A., Spinning test particles in general relativity II, Proc. Roy. Soc. Lond. A 209, 259-269 (1951).

[10] Dirac, P.A.M., Lectures on Quantum Mechanics, Belfer Graduate School of Science, Yeshiva University, New York, 1964.

[11] Dirac, P.A.M., Generalized Hamiltonian dynamics, Can. J. Math. 2, 129-148 (1950).

[12] Dixon, W.G., A covariant multipole formalism for extended test bodies in general relativity, Nuovo Cim. 34, 317-339 (1964).

[13] Dixon, W.G., Dynamics of extended bodies in general relativity I. Momentum and angular momentum, Proc. Roy. Soc. Lond. A 314, 499-527 (1970).

[14] Everitt, C.W.F. and Fairbank, W.M., Application of low-temperature techniques to a satellite test of general relativity, Proceedings of the Tenth International Conference on Low Temperature Physics, Moscow, 1966, Vol. 2B (Vinti, Moscow, 1967) 337-342.

[15] Godbillon, C., Géométrie Differential et Mécanique Analytique, Hermann, Paris, 1969.

[16] Goldstein, H., Classical Mechanics, Addison-Wesley, Reading, Mass., 1959.

[17] Hanson, A.J. and Regge, T., The relativistic spherical top, Annals of Physics 87, 498-566 (1974).

[18] Havas, P., Four-dimensional formulations of Newtonian mechanics and their relation to the special and the general theory of relativity, Rev. Mod. Phys. 36, 938-965 (1964).

[19] Hawking, S.W. and Ellis, G.F.R., The Large Scale Structure of Space-Time, Cambridge University Press, Cambridge, 1973.

[20] Hermann, Robert, Differential Geometry and the Calculus of Variations, Academic Press, New York, 1968.

[21] Hermann, Robert, Vector Bundles in Mathematical Physics, Vol. II, W.A. Benjamin Inc., 1970.

[22] Hermann, R., Geodesics and classical mechanics on Lie groups, J. Math. Phys. 13, 460-464 (1972).

[23] Klein, J., Formalisme Lagrangian et structure presque tangent, Société Mathématique de France, Journées Relativistes de Caen, 136-138 (1970).

[24] Kobayashi, S. and Nomizu, K., Foundations of Differential Geometry I, Wiley, New York, 1963.

[25] Krause, J., Some remarks on orthonormal tetrad transport, Int. J. Theor. Phys. 12, 35-45 (1975).

[26] Künzle, H.P., Degenerate Lagrangian systems, Ann. Inst. Henri Poincaré 11, 393-414 (1969).

[27] Künzle, H.P., Canonical dynamics of spinning particles in gravitational and electromagnetic fields, J. Math. Phys. 13, 739-745 (1972).

[28] Künzle, H.P., Galilei and Lorentz structures on space-time: Comparison of the corresponding geometry and physics, Ann. Inst. Henri Poincaré 17, 337-362 (1972).

[29] Künzle, H.P., Dynamics of a rigid test body in curved space time, Comm. Math. Phys. 27, 23-36 (1972).

[30] Lévy-Leblond, J.M., Galilei group and Galilean invariance, article in Group Theory and its Applications, Vol. II, Loebl (ed.), Academic Press, New York, 1971.

[31] Lichnerowicz, A., Variété symplectique et dynamique associée à une sous-variété, C.R. Acad. Sc. Paris 280, 523-527 (1975).

[32] MacLane, S. et al, Geometrical Mechanics Part I and Part II, Dept. Math., University of Chicago, 1968.

[33] Madore, J., Definition of centre of mass of an extended body in general relativity, Ann. Inst. Henri Poincaré 11, 221-237 (1969).

[34] Mercier, A., Analytical and Canonical Formalism in Physics, North-Holland, Amsterdam, 1959.

[35] Misner, C.W., Thorne, K.S. and Wheeler, J.A., Gravitation, W.H. Freeman and Co., San Fransisco, 1972.

[36] O'Connell, R.F., Present status of the relativity-gyroscope experiment, Gen. Rel. Grav. 3, 123-133 (1972).

[37] Palais, R., A Global Formulation of the Lie Theory of Transformation Groups, A.M.S. Memoir no. 22, 1957.

[38] Papapetrou, A., Spinning test-particles in general relativity. I, Proc. Roy. Soc. Lond. A 209, 248-258 (1951).

[39] Pauri, M. and Prosperi, G.M., Canonical realizations of the Galilei group, J. Math. Phys. 9, 1146-1162 (1968).

[40] Pinski, G., Galilean tensor calculus, J. Math. Phys. 9, 1927-1930 (1968).

[41] Pirani, F.A.E., Rigid motion in relativity theory, in Trautman (ed.), Lectures in General Relativity, Brandeis, 1964.

[42] Saletan, E.J. and Cromer, A.H., Theoretical Mechanics, John Wiley, New York, 1971.

[43] Schiff, L.I., Motion of a gyroscope according to Einstein's theory of gravitation, Proc. N.A.S., U.S. 46, 871-882 (1960).

[44] Slawianowski, J.J., The mechanics of an affinely rigid body, Int. J. Theor. Phys. 12, 271-296 (1975).

[45] Śniatycki, J., Geometrical theory of constraints, Proceedings, "Journées Relativiste de Caen 1970", 119-125.

[46] Śniatycki, J., Dirac brackets in geometrical dynamics, Ann. Inst. Henri Poincaré 20, 365-372 (1974).

[47] Śniatycki, J. and Tulczyjew, W.M., Canonical dynamics of relativistic charged particles, Ann. Inst. Henri Poincaré 15, 177-187 (1971).

[48] Souriau, J.M., Structure des Systèmes Dynamique, Dunod, Paris, 1970.

[49] Souriau, J.M., Modèle de particule à spin dans le champ électromagnétique et gravitationnel, Ann. Inst. Henri Poincaré 20, 315-364 (1974).

[50] Steenrod, N., The Topology of Fibre Bundles, Princeton University Press, Princeton, 1951.

[51] Sternberg, S., Lectures on Differential Geometry, Prentice-Hall, Englewood Cliffs; 1964.

[52] Sudarshan, E.C.G. and Mukunda, N., Classical Dynamics: A Modern Perspective, John Wiley, New York, 1974.

[53] Synge, J.L., Relativity: The General Theory, North-Holland, Amsterdam, 1966.

[54] Trautman, A., Theories of space, time and gravitation, in, Trautman (ed.), Lectures in General Relativity, Vol. I, Brandeis, 1964, Chapter 5.

[55] Trautman, A., Comparison of Newtonian and relativistic theories of space-time, article in, B. Hoffmann (ed.), Perspectives in Geometry and Relativity, Indiana University Press, Bloomington, 1966, 413-425.

[56] Trautman, A., The Einstein-Cartan theory of gravitation, Summary of a report prepared for the colloquium "Ondes et Radiations Gravitationnelles", Paris, 1973.

[57] Truesdell, C., The Elements of Continuum Mechanics, Springer-Verlag, New York, 1966.

[58] Vershik, A.M. and Faddeev, L.D., Differential geometry and Lagrangian mechanics with constraints, Soviet Physics-Doklady 17, 34-36 (1972).

[59] Warner, F.W., Foundations of Differential Manifolds and Lie Groups, Scott, Foresman and Co. Glenview, Illinois, 1971.

[60] Whiston, G.S., Topics on space-time topology (III), Int. J. Theor. Phys. 12, 225-240 (1975).

[61] Hirsch, M.W., On imbedding differentiable manifolds in Euclidean space, Ann. Math., 73, 566-571 (1961).

[62] Mathisson, M., New Mechanics of Material Systems, Acta Phys. Pol. 6, 163 (1937).

APPENDIX

Principal Bundles, Frame Bundles and Connections

(a) Group actions. Let G be a Lie group and M a manifold. A left action of G on M is a C^∞ - map

(1) $$\phi : G\times M \to M \ , \ (a,x) \to ax \equiv \psi(a,x)$$

such that $(ba)x = b(ax)$ for all $a,b \in G$ and $x \in M$. A right action is similarly defined as $\psi : M\times G \to M$. Sometimes the notation

(2) $$\begin{cases} R_a(x) = \psi(x,a) = xa \\ \\ L_a(x) = \psi(a,x) = ax \end{cases}$$

is used.

We say that G acts effectively on M iff $R_a(x) = x$ for all $x \in M$ implies $a = id$. G acts freely on M iff there is an $x \in M$ such that $R_a(x) = x$ implies $a = id$.

(b) Principal fibre bundle. Let M be a manifold and G a Lie group. A principal fibre bundle over M with group G is a manifold P together with an action of G on P satisfying:

(i) G acts freely on P on the right:

$$R : P\times G \to P \quad , \quad (u,a) \to R_a(u) = ua \quad .$$

(ii) M is the quotient space of P by the equivalence relation induced by G, and the canonical projection $\pi : P \to M$ is smooth.

(iii) P is locally trivial; that is, about each $x \in M$ there is a neighbourhood U such that $\pi^{-1}(U)$ is isomorphic to $U \times G$ in the sense that there is a diffeomorphism $\psi : \pi^{-1}(U) \to U \times G$ such that $\psi(u) = (\pi(u), \phi(u))$, where $\phi : \pi^{-1}(U) \to G$ satisfies $\phi(ua) = (\phi(u))a$ for all $u \in \pi^{-1}(U)$ and $a \in G$.

Call P the <u>bundle space</u>, M the <u>base space</u>, G the <u>structure group</u>, π the <u>projection</u> and $\pi^{-1}(x)$ the <u>fibre</u> over $x \in M$. (Note that $\pi^{-1}(x) \cong G$ for all $x \in M$.)

If M is any manifold and G a Lie group, then trivially we have $P = M \times G$ with the action $P \times G \to P$, $((x,a),b) \to (x,ab)$ as an example of a principal fibre bundle.

(c) Fundamental vector fields. Let P be a principal bundle with group G and base M. The action ψ of G on P induces an isomorphism σ of the Lie algebra g of G into $\mathfrak{X}(P)$: If $A \in g$ then $\exp(tA)$ is a 1-parameter subgroup of G; at $u \in P$, define

$$\sigma(A)\Big|_u = \frac{d}{dt}\Big|_{t=0} \psi(u, \exp(tA)) \quad . \tag{3}$$

$\sigma(A)$ is variously denoted in the literature by A^* or $\underset{\sim}{A}$.
(Sometimes $\underset{\sim}{A}$ is called a "Killing vector field".)

The mapping $A \to A^*|_u$ of $g \to T_u P$ is an isomorphism of g with the tangent space at u to the fibre through u : $g \cong T_u \pi^{-1}(\pi(u))$.

(d) Connection on a principal fibre bundle. Let P be a principal bundle over M with group G . If $u \in P$, G_u denotes the subspace of $T_u P$ tangent to the fibre $\pi^{-1}(u)$. A <u>connection</u> Γ in P is a smooth assignment of a subspace Q_u of the tangent space to each $u \in P$ which satisfies:

(i) $T_u P = G_u \oplus Q_u$

(ii) $Q_{(ua)} = (R_a)_* Q_u$.

G_u is called the <u>vertical subspace</u> and Q_u the <u>horizontal subspace</u> of $T_u P$. A vector in G_u is called <u>vertical</u>, while one in Q_u is called <u>horizontal</u>.

(e) Connection form. Let Γ be a connection on P and denote by $v(X)$ (resp. $h(X)$) the vertical (resp. horizontal) component of X with respect to Γ . Define the g - <u>valued connection 1-form</u> ω on P as follows: if $X \in T_u P$ then

$$(4) \qquad \begin{cases} \omega(X) = A \quad \text{where } A \text{ satisfies} \\ A^*|_u = v(X) \quad . \end{cases}$$

A is necessarily unique (since $g \cong G_u$) and $\omega(X) = 0$ iff X is horizontal.

(f) Bundle of linear frames of a manifold. Let M be a n - dimensional manifold. A linear frame at $x \in M$ is an ordered basis $e = (e_1, \ldots, e_n)$ of T_xM. Denote by L_xM the set of all such frames of T_xM. Then $L(M) = \bigcup_{x \in M} L_xM$ is a principal fibre bundle over M with projection $\pi : (x,e) \to x$ and a right action of $GL(n)$ by: $((x,e),a) \to (x, e\cdot a)$ where $(e\cdot a)_i = e_j a^j{}_i$.

Let (x^α) be a local coordinate system on a neighbourhood U of $x \in M$. Then each frame e of T_xM can be uniquely expressed as

(5) $$e_i = e_i^\alpha \frac{\partial}{\partial x^\alpha} \quad , \quad (e_i^\alpha) \in GL(n) \quad .$$

We take, then, as a local coordinate system at $(x,e) \in P$ (x^α, e_a^α) on the neighbourhood $\pi^{-1}(U)$.

Occasionally, we view a frame e at x as a non-singular map $R^n \to T_xM$. The action of $GL(n)$ on $L(M)$ is then the composition $R^n \xrightarrow{a} R^n \xrightarrow{e} T_xM$ for $a \in GL(n)$ and $e \in L_xM$.

(g) The <u>canonical</u> R^n - <u>valued</u> <u>1-form</u> θ on $L(M)$ is defined by

(6) $$\theta(X) = e^{-1}(\pi_*(X))$$

for $X \in T_uL(M)$, $u = (x,e)$ and e is interpreted as in the last paragraph. In a local coordinate system (x^α, e_a^α), we can write:

$$\text{(7)} \qquad \begin{cases} X = X^{\alpha} \dfrac{\partial}{\partial x^{\alpha}} + X^{\alpha}_{a} \dfrac{\partial}{\partial e^{\alpha}_{a}} \\[2ex] \pi_{*}(X) = X^{\alpha} \dfrac{\partial}{\partial x^{\alpha}} \\[2ex] \theta^{a}_{\alpha} \equiv (e^{-1})^{a}_{\alpha} \\[2ex] \theta = (\theta^{1},\ldots,\theta^{n}) \\[2ex] \theta^{a} = \theta^{a}_{\alpha}\, dx^{\alpha} \quad . \end{cases}$$

(h) A connection Γ on $L(M)$ is called a <u>linear connection on</u> M. Let $\{E^{a}_{b}\}$ be the standard basis of $g\ell(n)$; then in terms of a local coordinate system $(x^{\alpha},e^{\alpha}_{a})$ on $L(M)$, the $g\ell(n)$ - valued connection 1-form ω has the expression

$$\text{(8)} \qquad \begin{cases} \omega = \omega^{a}_{\ b}\, E^{b}_{a} \\[2ex] \omega^{a}_{\ b} = \theta^{a}_{\alpha}(d\, e^{\alpha}_{b} + \Gamma^{\alpha}_{\beta\gamma}\, e^{\gamma}_{b}\, dx^{\beta}) \\[2ex] \Gamma^{\alpha}_{\beta\gamma} \text{ depends on } x = \pi(x,e) \text{ only.} \end{cases}$$

(i) Standard horizontal vector fields. Given a linear connection on M and a vector $\xi \in R^{n}$, define the horizontal vector field $\underset{\sim}{\xi}$ by

$$\text{(9)} \qquad \begin{cases} \pi_{*}(\underset{\sim}{\xi}\big|_{u}) = e(\xi) \\[3ex] \xi \in R^{n} \text{ , } u = (x,e) \in L(M) \text{ , } e : R^{n} \to T_{x}M \quad . \end{cases}$$

$\underset{\sim}{\xi}$ is called the <u>standard horizontal vector field corresponding to</u>

$\xi \in R^n$. In a local coordinate system (x^α, e^α_a) ,

$$\underset{\sim}{\xi} = \xi^r e^\alpha_r \left(\frac{\partial}{\partial x^\alpha} - \Gamma^\rho_{\alpha\sigma} e^\sigma_a \frac{\partial}{\partial e^\rho_a}\right) .$$

The standard horizontal vector fields corresponding to the standard basis (δ_a) of R^n $((\delta_a)^b \equiv \delta^b_a)$,

$$(10) \qquad \underset{\sim}{e}_r = e^\alpha_r \left(\frac{\partial}{\partial x^\alpha} - \Gamma^\rho_{\alpha\sigma} e^\sigma_a \frac{\partial}{\partial e^\rho_a}\right) ,$$

forms a basis of the horizontal subspace Q_u for each $u = (x,e) \in L(M)$.

(j) Parallelism of $TL(M)$. Recall that to each $A \in g\ell(n)$ there is associated the fundamental vector field $\underset{\sim}{A}$ which spans the vertical subspace G_u at each $u \in L(M)$. Locally, we have

$$(11) \qquad \underset{\sim}{E}^a_b = e^\alpha_b \frac{\partial}{\partial e^\alpha_a} .$$

Thus, the $n+n^2$ vectors $\{\underset{\sim}{e}_a|_u, \underset{\sim}{E}^a_b|_u\}$ forms a basis of $T_uL(M)$ for each $u \in L(M)$: if $X \in T_uL(M)$, then

$$(12) \qquad X = v^a \underset{\sim}{e}_a + U^a_b \underset{\sim}{E}^b_a$$

determines local coordinates $(x^\alpha, e^\alpha_a; v^a, U^a_b)$ for $TL(M)$. Furthermore, the following orthonormality conditions hold:

$$(13) \qquad \begin{cases} \theta^a(\underset{\sim}{e}_b) = \delta^a_b & \theta^a(\underset{\sim}{E}^b_c) = 0 \\ \omega^a{}_b(\underset{\sim}{e}_c) = 0 & \omega^a{}_b(\underset{\sim}{E}^c_d) = \delta^a_d\, \delta^c_b \end{cases} .$$

(k) Structure equations of a linear connection. Define a 2-form Ω on $L(M)$ by:

$$\Omega|_u(X_1,X_2) = d\omega|_u(h(X_1),h(X_2)) \tag{14}$$

where $X_1,X_2 \in T_uL(M)$ and $h(\ )$ is the horizontal part. $\Omega = \Omega^a_b\, E^b_a$ is called the curvature 2-form of the connection and satisfies the structure equation

$$d\omega^a{}_b + \omega^a{}_r \wedge \omega^r{}_b = \Omega^a_b \ . \tag{15}$$

The canonical 1-form θ (equation (7)) satisfies a second structure equation:

$$d\theta^a + \omega^a{}_r \wedge \theta^r = \Theta^a \tag{16}$$

where $\Theta(X_1,X_2) = d\theta(h(X_1),h(X_2))$ is the torsion form of Γ . In this thesis, we assume $\Theta = 0$. (But see Trautman [54] or [56]).